Lower Bounds for Unambiguous Automata via Communication Complexity

Mika Göös
EPFL, Lausanne, Switzerland

Stefan Kiefer
University of Oxford, UK

Weiqiang Yuan
EPFL, Lausanne, Switzerland

Abstract

We use results from communication complexity, both new and old ones, to prove lower bounds for unambiguous finite automata (UFAs). We show three results.

1. *Complement:* There is a language L recognised by an n-state UFA such that the complement language $\overline{L}$ requires NFAs with $n^{\tilde{\Omega}(\log n)}$ states. This improves on a lower bound by Raskin.
2. *Union:* There are languages L_1, L_2 recognised by n-state UFAs such that the union $L_1 \cup L_2$ requires UFAs with $n^{\tilde{\Omega}(\log n)}$ states.
3. *Separation:* There is a language L such that both L and $\overline{L}$ are recognised by n-state NFAs but such that L requires UFAs with $n^{\Omega(\log n)}$ states. This refutes a conjecture by Colcombet.

2012 ACM Subject Classification Theory of computation → Regular languages

Keywords and phrases Unambiguous automata, communication complexity

Digital Object Identifier 10.4230/LIPIcs.ICALP.2022.126

Category Track B: Automata, Logic, Semantics, and Theory of Programming

Acknowledgements We thank anonymous ICALP reviewers for many comments.

1 Introduction

Given two finite automata recognising languages $L_1, L_2 \subseteq \Sigma^*$ a basic question is to determine the *state complexity* of various language operations. How many states are needed in an automaton that recognises the union $L_1 \cup L_2$? How about the intersection $L_1 \cap L_2$? The complement $\overline{L}_1 := \Sigma^* \setminus L_1$? The answer depends on the type of automaton considered, such as deterministic (DFA), nondeterministic (NFA), or unambiguous (UFA). Recall that a UFA is an NFA that has at most one accepting computation on any input.

State complexities have been extensively studied for various types of automata and language operations; see, e.g., [9, 15] and their references, or the excellent compendium on Wikipedia [22]. For example, complementing an NFA with n states may require 2^n states [3], even for automata with binary alphabet [14]. Surprisingly, several extremely basic questions about UFAs remain open. For example, it was shown only in 2018 by Raskin [20] that the state complexity for UFA complementation is not polynomial: for any $n \in \mathbb{N}$ there exists a language L recognised by an n-state UFA such that any UFA (or even NFA) that recognises $\overline{L}$ has at least $n^{(\log \log \log n)^{\Omega(1)}}$ states. This superpolynomial blowup refuted a conjecture that it may be possible to complement UFAs with a polynomial blowup [5].

In this paper, as our main results, we prove three new blowup theorems.

▶ **Theorem 1** (Complement). *For every $n \in \mathbb{N}$ there is a language $L \subseteq \{0,1\}^*$ recognised by an n-state UFA such that any NFA that recognises $\overline{L}$ requires $n^{\tilde{\Omega}(\log n)}$ states.*

▶ **Theorem 2** (Union). *For every $n \in \mathbb{N}$ there are languages $L_1, L_2 \subseteq \{0,1\}^*$ recognised by n-state UFAs such that any UFA that recognises $L_1 \cup L_2$ requires $n^{\tilde{\Omega}(\log n)}$ states.*

49th International Colloquium on Automata, Languages, and Programming (ICALP 2022).
Editors: Mikołaj Bojańczyk, Emanuela Merelli, and David P. Woodruff;
Article No. 126; pp. 126:1–126:13
Leibniz International Proceedings in Informatics
Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

▶ **Theorem 3** (Separation). *For every $n \in \mathbb{N}$ there is a language $L \subseteq \{0,1\}^*$ such that both L and $\overline{L}$ are recognised by n-state NFAs but any UFA that recognises L requires $n^{\Omega(\log n)}$ states.*

Discussion of main results. **Theorem 1** upgrades Raskin's slightly-superpolynomial bound into a quasipolynomial bound $n^{\tilde{\Omega}(\log n)}$. (Here we use the notation $\tilde{\Omega}(m)$ to suppress $\mathrm{poly}(\log m)$ factors.) However, we note that Raskin's language is unary, $|\Sigma| = 1$, while ours is binary, $|\Sigma| = 2$, and hence the two results are incomparable in this sense. As for positive results, it is known that the trivial 2^n upper bound for UFA complementation can be improved: the complement of any n-state UFA can be recognised by a UFA with at most $\mathrm{poly}(n) \cdot 2^{n/2}$ states [15, 13]. Closing the exponential gap here between the lower and upper bounds remains a tantalising open problem. It was highlighted as one of the foremost challenges in the recent Dagstuhl workshop *Unambiguity in Automata Theory* [6].

Theorem 2 establishes the first superpolynomial lower bound for the union operation. Letting $\sqcup$ denote disjoint union, observe that

$$L_1 \cup L_2 \;=\; L_1 \sqcup (L_2 \cap \overline{L}_1). \tag{1}$$

Since disjoint union and intersection are polynomial for UFAs, it follows from (1) and Theorem 2 that the same $n^{\tilde{\Omega}(\log n)}$ lower bound holds for complementing UFAs. However, we stress that Theorem 1 has a stronger conclusion than this, since it proves a lower bound against NFAs, not just UFAs. The observation (1) also yields the upper bound $\mathrm{poly}(n) \cdot 2^{n/2}$ by using the complement construction from [15, 13].

Theorem 3 refutes a conjecture by Colcombet [5, Conjecture 2]. Indeed, he conjectured that for any pair of NFAs recognising languages L_1, L_2 such that $L_1 \cap L_2 = \emptyset$, there is a polynomial-sized UFA that recognises some L that *separates* L_1 and L_2 in the sense that $L_1 \subseteq L$ and $L \cap L_2 = \emptyset$. Theorem 3 refutes this even in the special case $L_1 = \overline{L}_2$. Related separability questions are classical in formal language theory and have attracted renewed attention; see, e.g, [8] and the references therein. Separating automata have also been used recently to elegantly describe quasipolynomial time algorithms for solving parity games in an automata theoretic framework; see [4, Chapter 3] and [7].

1.1 Technique: Communication complexity

Our three main theorems rely on results – both new and old – in communication complexity; see [18, 19] for the standard textbooks. In communication complexity, one studies functions of the form $F\colon \{0,1\}^n \times \{0,1\}^n \to \{0,1\}$ that determine the following two-party communication problem: Alice holds $x \in \{0,1\}^n$, Bob holds $y \in \{0,1\}^n$, and their goal is to output $F(x,y)$ while communicating as few bits as possible between them. Communication complexity is a classical tool to prove lower bounds for automata. Indeed, it is well known that if the language $\{xy : F(x,y) = 1\}$ is recognised by a small DFA (resp. NFA, UFA) then F admits an efficient deterministic (resp. nondeterministic, unambiguous) protocol. We revisit this connection in light of recent developments in communication complexity.

Theorem 1 is a ***relatively straightforward consequence*** of a recent result of Balodis et al. [2]. They exhibited a two-party function whose co-nondeterministic communication complexity is nearly quadratic in its unambiguous complexity (which matches an upper bound due to Yannakakis [23]). We translate this separation into the language of automata theory, virtually in a black-box fashion.

Theorem 2, by contrast, is our ***main technical contribution***. We will show that it follows from the following analogous communication result, which we prove in this paper.

▶ **Theorem 4.** *For every $m \in \mathbb{N}$ there exists a function $F(x, y)$ with unambiguous communication complexity at most m such that the logical-or of two copies of F, namely, $F^{\vee}(xx', yy') := F(x, y) \vee F(x', y')$, has unambiguous communication complexity $\tilde{\Omega}(m^2)$.*

This is a new result in communication complexity; the unambiguous complexity of $F^{\vee}$ has not been studied previously. We prove Theorem 4 using the popular *query-to-communication lifting* technique that has been wildly successful in the past decade to prove communication lower bounds (including in [2]). In this technique, one starts by proving a lower bound on the query (aka decision tree) complexity of a boolean function $f\colon \{0,1\}^n \to \{0,1\}$. A lifting theorem (e.g., [12]) then transforms f into an analogous communication problem F in such a way that the communication complexity of F is characterised by the query complexity of f. This reduces the task of proving communication lower bounds into the much easier task of proving query lower bounds.

Interestingly, our proof of Theorem 4 formalises a kind of *converse* to the observation (1) above (saying that union can be computed via a complement). Namely, we show that unambiguously computing the union necessarily requires computing a complement, and therefore we can rely on an existing query lower bound for complementation [11].

Theorem 3, finally, is a straightforward consequence of a classical quadratic separation between two-sided nondeterministic communication complexity and unambiguous communication complexity due to Razborov [21].

1.2 Bonus result: Approximate nonnegative rank

Along the way to Theorem 4 we inadvertently stumbled upon another separation result that addresses a question raised by Kol et al. [16]. They studied the *ϵ-approximate nonnegative rank* $\mathrm{rk}^+_\epsilon(M)$ of a nonnegative matrix $M \in \mathbb{R}^{n\times n}$. Here, $\mathrm{rk}^+_\epsilon(M)$ is defined as the least nonnegative rank $\mathrm{rk}^+(N)$ of a matrix $N \in \mathbb{R}^{n\times n}$ such that $|M_{ij} - N_{ij}| \le \epsilon$ for all i, j; see Section 3 for precise definitions. In particular, Kol et al. [16] asked whether for all error parameters $0 < \epsilon < \delta < 1/2$ and boolean matrices $M \in \{0,1\}^{n\times n}$ we have the polynomial relationship $\mathrm{rk}^+_\epsilon(M) \le O(\mathrm{rk}^+_\delta(M)^C)$ where $C = C(\epsilon, \delta)$ is a constant. In short, does approximate nonnegative rank admit efficient error reduction? (It is known that the more usual notion, *approximate rank*, does [1].) We provide the following negative answer.

▶ **Theorem 5** (No efficient error reduction)**.** *For every $m \in \mathbb{N}$ there exists a boolean matrix M with $\mathrm{rk}^+_{1/4}(M) \le m$ but such that $\mathrm{rk}^+_{10^{-5}}(M) \ge m^{\tilde{\Omega}(\log m)}$.*

Previously, a negative answer was known only for *partial* boolean matrices $M \in \{0, 1, *\}^{n\times n}$ that allow “don’t care” entries $M_{ij} = *$ [12]. Our Theorem 5 still leaves open the possibility (also raised by [16]) that, for a total boolean matrix M, we can bound $\mathrm{rk}^+_\epsilon(M)$ as a polynomial function of $\mathrm{rk}^+_\delta(M) + \mathrm{rk}^+_\delta(\overline{M})$ where $\overline{M}$ is the boolean complement.

1.3 Open problems

Our quasipolynomial lower bounds for automata are not known to be tight; in all cases the best known upper bounds are exponential. Curiously enough, the analogous communication results are tight for communication protocols. This suggests two opportunities.

- Can other techniques from communication complexity improve the lower bounds further? Perhaps via multi-party communication complexity?
- Can techniques for proving upper bounds on communication complexity be adapted to prove upper bounds on the size of automata?

1.4 Definitions of automata

An *NFA* is a quintuple $\mathcal{A} = (Q, \Sigma, \delta, I, F)$, where Q is the finite set of states, Σ is the finite alphabet, $\delta \subseteq Q \times \Sigma \times Q$ is the transition relation, $I \subseteq Q$ is the set of initial states, and $F \subseteq Q$ is the set of accepting states. We write $q \xrightarrow{a} r$ to denote that $(q, a, r) \in \delta$. A finite sequence $q_0 \xrightarrow{a_1} q_1 \xrightarrow{a_2} \cdots \xrightarrow{a_n} q_n$ is called a *run*; it can be summarized as $q_0 \xrightarrow{a_1 \cdots a_n} q_n$. The NFA $\mathcal{A}$ *recognizes* the language $L(\mathcal{A}) := \{w \in \Sigma^* \mid \exists\, q_0 \in I \,.\, \exists\, f \in F \,.\, q_0 \xrightarrow{w} f\}$. The NFA $\mathcal{A}$ is a *DFA* if $|I| = 1$ and for every $q \in Q$ and $a \in \Sigma$ there is exactly one q' with $q \xrightarrow{a} q'$. The NFA $\mathcal{A}$ is a *UFA* if for every word $w = a_1 \cdots a_n \in \Sigma^*$ there is at most one *accepting* run for w, i.e., a run $q_0 \xrightarrow{a_1} q_1 \xrightarrow{a_2} \cdots \xrightarrow{a_n} q_n$ with $q_0 \in I$ and $q_n \in F$. Any DFA is a UFA.

2 UFA Complementation

In this section we prove Theorem 1.

▶ **Theorem 1** (Complement). *For every $n \in \mathbb{N}$ there is a language $L \subseteq \{0,1\}^*$ recognised by an n-state UFA such that any NFA that recognises $\overline{L}$ requires $n^{\tilde{\Omega}(\log n)}$ states.*

The proof uses concepts from communication complexity, in particular a recent result from [2] and a nondeterministic lifting theorem from [12]. We start by recalling these tools.

2.1 DNFs and nondeterministic protocols

Unambiguous DNFs. Let $D = C_1 \vee \cdots \vee C_m$ be an n-variate boolean formula in disjunctive normal form (DNF). DNF D has *width* k if every C_i is a conjunction of at most k literals. We call such D a *k-DNF*. For conjunctive normal form (CNF) formulas the width and k-CNFs are defined analogously. DNF D is said to be *unambiguous* if for every input $x \in \{0,1\}^n$ at most one of the conjunctions C_i evaluates to true, $C_i(x) = 1$. For any boolean function $f\colon \{0,1\}^n \to \{0,1\}$ define

- $\mathrm{C}_1(f)$ as the least k such that f can be written as a k-DNF;
- $\mathrm{C}_0(f)$ as the least k such that f can be written as a k-CNF;
- $\mathrm{UC}_1(f)$ as the least k such that f can be written as an unambiguous k-DNF.

Note that $\mathrm{C}_0(f) = \mathrm{C}_1(\neg f)$. The following recent result separates two of these measures.

▶ **Theorem 6** ([2, Theorem 1]). *For every $k \in \mathbb{N}$ there exists a function $f\colon \{0,1\}^n \to \{0,1\}$ where $n \le \mathrm{poly}(k)$ and such that $\mathrm{UC}_1(f) \le k$ and $\mathrm{C}_0(f) \ge \tilde{\Omega}(k^2)$.* ◀

In words, for every k there is an unambiguous k-DNF such that any equivalent CNF requires width $\tilde{\Omega}(k^2)$. The bound is almost tight, as every unambiguous k-DNF has an equivalent k^2-CNF; see [10, Section 3].

Nondeterministic protocols and rectangle covers. Next we recall standard notions from two-party communication complexity; see [18, 19] for textbooks. Consider a two-party function $F\colon X \times Y \to \{0,1\}$. A set $A \times B \subseteq X \times Y$ (with $A \subseteq X$ and $B \subseteq Y$) is called a *rectangle*. Rectangles $R_1, \ldots, R_k$ *cover* a set $S \subseteq X \times Y$ if $\bigcup_i R_i = S$. For $b \in \{0,1\}$, the *cover number*

$\mathrm{Cov}_b(F)$ is the least number of rectangles that cover $F^{-1}(b)$. The *nondeterministic (resp., co-nondeterministic) communication complexity* of F is defined as $\mathrm{N}_1(F) \coloneqq \log_2 \mathrm{Cov}_1(F)$ (resp., $\mathrm{N}_0(F) \coloneqq \log_2 \mathrm{Cov}_0(F)$). Note that $\mathrm{N}_0(F) = \mathrm{N}_1(\neg F)$. The nondeterministic communication complexity can be interpreted as the number of bits that two parties (Alice and Bob), holding inputs $x \in X$ and $y \in Y$, respectively, need to communicate in a nondeterministic (i.e., based on guessing and checking) protocol in order to establish that $F(x, y) = 1$; see [18, Chapter 2] for details.

Nondeterministic lifting. Next we formulate a *lifting theorem*, which allows us to transfer lower bounds on the DNF width of an n-bit boolean function f to the nondeterministic communication complexity of a related two-party function F. We first choose a small two-party function $g\colon \{0,1\}^b \times \{0,1\}^b \to \{0,1\}$, often called a *gadget*. Then we compose f with g to construct the function $F \coloneqq f \circ g^n$ that maps $\{0,1\}^{bn} \times \{0,1\}^{bn} \to \{0,1\}$ where Alice gets as input $x \in \{0,1\}^{bn}$, Bob gets as input $y \in \{0,1\}^{bn}$, and their goal is to compute

$$F(x, y) \;\coloneqq\; f(g(x_1, y_1), \ldots, g(x_n, y_n)) \qquad \text{where } x_i, y_j \in \{0,1\}^b.$$

The following is a nondeterministic lifting theorem [12, 10].

▶ **Theorem 7** ([10, Theorem 4])**.** *For any $n \in \mathbb{N}$ there is a gadget $g\colon \{0,1\}^b \times \{0,1\}^b \to \{0,1\}$ with $b = \Theta(\log n)$ such that for any function $f\colon \{0,1\}^n \to \{0,1\}$ we have, for $F \coloneqq f \circ g^n$,*

$$\mathrm{N}_0(F) \;=\; \Omega(\mathrm{C}_0(f) \cdot b)$$

(and thus also $\mathrm{N}_1(F) = \Omega(\mathrm{C}_1(f) \cdot b)$). ◀

Protocols can simulate automata. Finally, we need a simple folklore connection between automata and protocols. To formalise this, we tacitly identify a function $F\colon \{0,1\}^{m_1} \times \{0,1\}^{m_2} \to \{0,1\}$ with the language $F^{-1}(1) \;=\; \{xy \in \{0,1\}^{m_1+m_2} \mid F(x, y) = 1\}$.

▶ **Lemma 8.** *If a two-party function $F\colon \{0,1\}^{m_1} \times \{0,1\}^{m_2} \to \{0,1\}$ admits an NFA with s states, then $\mathrm{Cov}_1(F) \le s$ (that is, $\mathrm{N}_1(F) \le \log s$).*

Proof. Let $\mathcal{A} = (Q, \Sigma, \delta, I, F)$ be an NFA with $L(\mathcal{A}) = \{xy \in \{0,1\}^{m_1+m_2} \mid F(x, y) = 1\}$. We show that $F^{-1}(1)$ is covered by at most $|Q|$ rectangles. Indeed, $F^{-1}(1)$ equals

$$\bigcup_{q \in Q} (\{x \in \{0,1\}^{m_1} \mid \exists\, q_0 \in I \,.\, q_0 \xrightarrow{x} q\}) \times (\{y \in \{0,1\}^{m_2} \mid \exists\, f \in F \,.\, q \xrightarrow{y} f\})\,.$$

(Alternatively, in terms of a nondeterministic protocol, the first party, holding $x \in \{0,1\}^{m_1}$, produces a run for x from an initial state to a state q and then sends the name of q, which takes $\log_2 |Q|$ bits, to the other party. The other party then produces a run for y from q to an accepting state.) ◀

2.2 Proof of Theorem 1

For $k \in \mathbb{N}$, let $f\colon \{0,1\}^n \to \{0,1\}$ be the function from Theorem 6. That is, f has an unambiguous k-DNF with $k = n^{\Omega(1)}$ (hence, $\log n = O(\log k)$) and $\mathrm{C}_0(f) = \tilde{\Omega}(k^2)$. Let $g\colon \{0,1\}^b \times \{0,1\}^b \to \{0,1\}$ with $b = \Theta(\log n)$ and $F \coloneqq f \circ g^n \colon \{0,1\}^{bn} \times \{0,1\}^{bn} \to \{0,1\}$ be the two-party functions from the lifting theorem Theorem 7. We will show that Theorem 1 holds for the language $F^{-1}(1)$.

First we argue that F has an unambiguous DNF of small width. Indeed, g and $\neg g$ have unambiguous $2b$-DNFs, which can be extracted from the deterministic decision tree of g. By plugging these unambiguous $2b$-DNFs for g and $\neg g$ into the unambiguous k-DNF for f (and "multiplying out"), one obtains an unambiguous $2bk$-DNF, say D, for F.

Over the $2bn$ variables of F, there exist at most $(2(2bn)+1)^{2bk}$ different conjunctions of at most $2bk$ literals. So D consists of at most $n^{O(bk)}$ conjunctions. From D we obtain a UFA $\mathcal{A}$ that recognizes $F^{-1}(1) \subseteq \{0,1\}^{2bn}$, as follows. Each initial state of $\mathcal{A}$ corresponds to a conjunction in D. When reading the input $x \in \{0,1\}^{2bn}$, the UFA checks that the corresponding assignment to the variables satisfies the conjunction represented by the initial state. This check requires at most $O(bn)$ states for each initial state. Thus, $\mathcal{A}$ has at most $n^{O(bk)} = 2^{\tilde{O}(k)} =: N$ states in total. (We use N in place of n in the statement of Theorem 1.)

On the other hand, by Theorem 7, we have $\mathrm{N}_0(F) = \Omega(\mathrm{C}_0(f) \cdot b) = \tilde{\Omega}(k^2)$. So by Lemma 8 any NFA that recognizes $F^{-1}(0)$ has at least $2^{\tilde{\Omega}(k^2)}$ states. Any NFA that recognizes $\{0,1\}^* \setminus L(\mathcal{A})$ can be transformed into an NFA that recognizes $F^{-1}(0) = \{0,1\}^{2bn} \setminus L(\mathcal{A})$ by taking a product with a DFA that has $2bn+2$ states. It follows that any NFA that recognizes $\{0,1\}^* \setminus L(\mathcal{A})$ has at least $2^{\tilde{\Omega}(k^2)}/(2bn+2) = 2^{\tilde{\Omega}(k^2)} = N^{\tilde{\Omega}(\log N)}$ states. ◀

3 UFA Union

In this section, we prove Theorem 2.

▶ **Theorem 2** (Union). *For every $n \in \mathbb{N}$ there are languages $L_1, L_2 \subseteq \{0,1\}^*$ recognised by n-state UFAs such that any UFA that recognises $L_1 \cup L_2$ requires $n^{\tilde{\Omega}(\log n)}$ states.*

We follow the same high-level approach that we already saw in Section 2. Namely, we will first show that computing the $\vee$-operation is hard for unambiguous DNFs and then lift that hardness to unambiguous protocols, which then implies the same hardness for UFAs. There are, however, two challenges in carrying out this plan.

1. It is an open problem to prove an unambiguous lifting theorem. That is, it is not known whether the unambiguous communication complexity of $f \circ g^n$ is at least $\Omega(\mathrm{UC}_1(f))$. To circumvent this issue, we study instead a *linear relaxation* of unambiguous DNFs. These objects are called *conical juntas* and they do admit a lifting theorem [12, 17].
2. There is no existing result showing that the $\vee$-operation is hard for unambiguous DNFs and/or conical juntas. We show a result of this type. The proof is by a reduction to the hardness of negating conical juntas, which *is* a known result [11].

3.1 Conical juntas

A nonnegative function $h\colon \{0,1\}^n \to \mathbb{R}_{\geq 0}$ is a *d-junta* if h depends on at most d variables. For example, a conjunction of d literals is a d-junta. Moreover, we say $f\colon \{0,1\}^n \to \mathbb{R}_{\geq 0}$ is a *conical d-junta* if it can be written as a nonnegative linear combination of d-juntas. Equivalently, f is a conical d-junta if it can be written as $f = \sum_i w_i C_i$ where each C_i is a width-d conjunction and $w_i \in \mathbb{R}_{\geq 0}$ are nonnegative coefficients. For example, if f can be written as an unambiguous d-DNF, $f = C_1 \vee \cdots \vee C_m$, then $f = \sum_i C_i$ is a conical d-junta with 0/1 coefficients. The *nonnegative degree* of f, denoted $\deg^+(f)$, is the least d such that f is a conical d-junta. In particular, if f is boolean-valued, then $\deg^+(f) \leq \mathrm{UC}_1(f)$.

We also need to work with *approximate* conical juntas that compute a given function only to within some point-wise error $\epsilon > 0$. This is important because the available lifting theorems for $\deg^+$ incur some error, and hence we need to prove lower bounds that are robust to this error. Indeed, we define the *ϵ-approximate nonnegative degree* of f, denoted $\deg^+_\epsilon(f)$, as the least nonnegative degree of a conical junta g such that

$$|f(x) - g(x)| \;\leq\; \epsilon \qquad \text{for all } x \in \{0,1\}^n.$$

▶ Remark. An awkward aspect of working with approximate conical juntas is that the error parameter ϵ is not well behaved. For $0 < \epsilon < \delta < 1/2$ we of course have $\deg^+_\delta(f) \leq \deg^+_\epsilon(f)$ but it is not a priori clear whether the converse inequality holds with a modest loss in the degree. In fact, in Section 5, we will end up showing that there can be a polynomial gap between the nonnegative degrees corresponding to two different error parameters – and this is related to our bonus result discussed in the introduction. As a consequence, our theorems in this section have to track the error parameters with some care.

Linear programming formulation. Approximate nonnegative degree can be captured using an LP. Write $\mathcal{C}^n_d$ for the set of all conjunctions of width at most d over n variables. In the (Primal) programme below, we have a variable $w_C \in \mathbb{R}$ for every $C \in \mathcal{C}^n_d$. In the associated (Dual) programme, we have a variable $\Phi(x) \in \mathbb{R}$ for each $x \in \{0,1\}^n$.

$$\begin{array}{rlll} \min & \epsilon & & \\ \textit{subject to} & \left|\sum_C w_C C(x) - f(x)\right| \;\leq\; \epsilon, & \forall x \in \{0,1\}^n & \text{(Primal)} \\ & w_C \;\geq\; 0, & \forall C \in \mathcal{C}^n_d & \end{array}$$

$$\begin{array}{rlll} \max & \langle \Phi, f\rangle \;\coloneqq\; \sum_x \Phi(x) f(x) & & \\ \textit{subject to} & \|\Phi\| \;\coloneqq\; \sum_x |\Phi(x)| \;\leq\; 1 & & \text{(Dual)} \\ & \langle \Phi, C\rangle \;\leq\; 0, & \forall C \in \mathcal{C}^n_d & \end{array}$$

We have that $\deg^+_\delta(f) \leq d$ iff the optimal value of (Primal) is at most δ. Alternatively, by strong LP duality, we have $\deg^+_\delta(f) > d$ iff there exists a feasible solution Φ to (Dual) such that $\langle \Phi, f\rangle > \delta$. It is typical to think of such feasible $\Phi\colon \{0,1\}^n \to \mathbb{R}$ as a *dual certificate* that witnesses a lower bound on approximate nonnegative degree.

3.2 Hardness of $\vee$

The goal of this subsection is to prove Theorem 9 below, which states that the $\vee$-operation is hard for unambiguous DNFs and even approximate conical juntas. Given an n-bit boolean function f we define a $2n$-bit function by $f^\vee(xy) \coloneqq f(x) \vee f(y)$ where $x, y \in \{0,1\}^n$.

▶ **Theorem 9** (Hardness of $\vee$). *For every $m \in \mathbb{N}$, there exists a boolean function $f\colon \{0,1\}^n \to \{0,1\}$ with $n \leq \mathrm{poly}(m)$ such that $\mathrm{UC}_1(f) \leq m$ and $\deg^+_{1.5\times 10^{-5}}(f^\vee) \geq \tilde{\Omega}(m^2)$.*

We show Theorem 9 by combining two lemmas, Lemmas 10 and 11, below. The first lemma, proved in [11], states that unambiguous DNFs are hard to negate, even by approximate conical juntas. The second lemma, which remains for us to prove, states that, for conical juntas, computing $f^\vee$ is at least as hard as computing the negation $\neg f$. Hence Theorem 9 follows immediately by combining these lemmas.

▶ **Lemma 10** (Hardness of ¬ [11, Lemma 8]). *For every $m \in \mathbb{N}$, there exists a boolean function $f\colon \{0,1\}^n \to \{0,1\}$ with $n \le \mathrm{poly}(m)$ such that $\mathrm{UC}_1(f) \le m$ and $\deg^+_{0.05}(\neg f) \ge \tilde{\Omega}(m^2)$.* ◀

▶ **Lemma 11** (∨ harder than ¬). *For every $\delta > 0$ there exists an $\epsilon = \epsilon(\delta) > 0$ such that for every boolean function f, we have $\deg^+_\epsilon(f^\vee) \ge \Omega(\deg^+_\delta(\neg f))$. Moreover, $\epsilon := \big(\frac{\ln(1+\delta)}{\lceil \log_{3/4} \delta \rceil}\big)^2$.*

It remains to prove Lemma 11. We do it in two steps. In Claim 12 we show that the approximate nonnegative degree of $f^\vee$ is at least that of $2 - f$ by exhibiting a dual certificate. Then in Claim 13 we show that the approximate nonnegative degree of $2 - f = 1 + \neg f$ is at least that of $\neg f$ via a powering trick. The error parameter ϵ will degrade in both of these steps. (We will later see that this degradation is, in fact, unavoidable; see Section 5.)

▷ Claim 12. We have $\deg^+_{\epsilon^2}(f^\vee) \ge \deg^+_\epsilon(2 - f)$ for any boolean-valued f and error ϵ.

Proof. Let $d := \deg^+_\epsilon(2 - f)$ and let $\Phi\colon \{0,1\}^n \to \{0,1\}$ be a dual certificate witnessing this. That is, $\langle \Phi, 2 - f\rangle > \epsilon$, $\|\Phi\| \le 1$, and $\langle \Phi, C\rangle \le 0$ for all $C \in \mathcal{C}^n_{d-1}$. To construct a dual certificate $\Phi^\vee\colon \{0,1\}^{2n} \to \{0,1\}$ witnessing $\deg^+_{\epsilon^2}(f^\vee) \ge d$, we consider the negated tensor product (which was found by an educated guess)

$$\Phi^\vee(x, y) \;:=\; -\Phi(x)\Phi(y).$$

It remains to check that this is feasible for the dual programme and also that $\langle \Phi^\vee, f^\vee\rangle > \epsilon^2$.

1. $\|\Phi^\vee\| = \sum_{x,y} |\Phi^\vee(x, y)| = \sum_{x,y} |\Phi(x)\Phi(y)| = \sum_{x,y} |\Phi(x)| \cdot |\Phi(y)| = \|\Phi\|^2 \le 1$.
2. For any conjunction $C \in \mathcal{C}^{2n}_{d-1}$, we write $C(x, y) = C_1(x)C_2(y)$ where $C_1, C_2 \in \mathcal{C}^n_{d-1}$. Now

$$\begin{aligned}
\langle \Phi^\vee, C\rangle &= \textstyle\sum_{x,y} \Phi^\vee(x, y)C(x, y)\\
&= \textstyle\sum_{x,y} -\Phi(x)\Phi(y) \cdot C_1(x)C_2(y)\\
&= -\big[\textstyle\sum_x \Phi(x)C_1(x)\big]\big[\textstyle\sum_y \Phi(y)C_2(y)\big]\\
&= -\langle \Phi, C_1\rangle \langle \Phi, C_2\rangle\\
&\le 0.
\end{aligned}$$

3. Observe that $\langle \Phi, -f\rangle \ge \langle \Phi, 2 - f\rangle$ since $1 \in \mathcal{C}^n_{d-1}$ for the constant-1 function. Thus

$$\begin{aligned}
\langle \Phi^\vee, f^\vee\rangle &= \textstyle\sum_{x,y} \Phi^\vee(x, y)f^\vee(x, y)\\
&= \textstyle\sum_{x,y} -\Phi(x)\Phi(y)(f(x) + f(y) - f(x)f(y))\\
&= \textstyle\sum_{x,y} -\Phi(x)\Phi(y)(2f(x) - f(x)f(y))\\
&= \textstyle\sum_x -\Phi(x)f(x) \cdot \big[\textstyle\sum_y \Phi(y)(2 - f(y))\big]\\
&= \langle \Phi, -f\rangle \cdot \langle \Phi, 2 - f\rangle\\
&\ge \langle \Phi, 2 - f\rangle \cdot \langle \Phi, 2 - f\rangle && \text{(above observation)}\\
&> \epsilon^2.
\end{aligned}$$

where the third equality holds since we have $\sum_{x,y} \Phi(x)\Phi(y)f(x) = \sum_{x,y} \Phi(x)\Phi(y)f(y)$ by exchanging x and y. ◁

▷ Claim 13. For any $\delta > 0$ define $\epsilon := \frac{\ln(1+\delta)}{\lceil \log_{3/4} \delta \rceil} > 0$. Then for any boolean-valued function f we have $\deg^+_\epsilon(1 + f) \ge \Omega(\deg^+_\delta(f))$.

Proof. We may assume $\delta < 1/2$ (and hence $\epsilon < 1/4$) as otherwise the claim is trivial. Suppose $\deg^+_\epsilon(1 + f) = d$ is witnessed by a conical d-junta g that ϵ-approximates $1 + f$. Define $g' := ((g + \epsilon)/2)^k$ where the exponent is $k := \lceil \log_{3/4} \delta \rceil$. By multiplying out the terms in this definition, we see that g' has nonnegative degree $kd = O(d)$. We claim that g' is a δ-approximation of f. Indeed, if $f(x) = 0$, then $g'(x) \leq (1/2 + \epsilon)^k \leq (3/4)^k \leq \delta$. If $f(x) = 1$, then $1 \leq (g(x) + \epsilon)/2 \leq 1 + \epsilon$, and thus $1 \leq g'(x) \leq (1 + \epsilon)^k \leq \exp(\epsilon k) \leq 1 + \delta$. ◁

Proof of Lemma 11. Using Claim 12 and Claim 13 (but with $\neg f$ in place of f), we have, for any $\delta > 0$ and $\epsilon := \frac{\ln(1+\delta)}{\lceil \log_{3/4} \delta \rceil}$:

$$\deg^+_{\epsilon^2}(f^\vee) \;\geq\; \deg^+_\epsilon(2 - f) \;=\; \deg^+_\epsilon(1 + \neg f) \;\geq\; \Omega(\deg^+_\delta(\neg f)).$$ ◀

3.3 Unambiguous protocols and nonnegative rank

Our goal will be to *lift* the hardness of the $\vee$-operation (Theorem 9) to communication complexity. In this subsection, we recall the concepts that are needed for this goal, namely, unambiguous protocols, (approximate) nonnegative rank, and a lifting theorem from nonnegative degree to nonnegative rank [12, 17].

Unambiguous protocols. Recall from Section 2.1 the notions of nondeterministic protocols and rectangle covers. For a two-party function $F\colon X \times Y \to \{0,1\}$, the *partition number* $\mathrm{Par}_1(F)$ is the least number of *pairwise disjoint* rectangles that cover $F^{-1}(1)$. Note that $\mathrm{Cov}_1(F) \leq \mathrm{Par}_1(F)$. The *unambiguous communication complexity* of F is defined as $\mathrm{U}_1(F) := \log_2 \mathrm{Par}_1(F)$. Note that $\mathrm{N}_1(F) \leq \mathrm{U}_1(F)$. Unambiguous communication complexity can be interpreted as the least communication cost of a nondeterministic protocol that has at most one accepting computation on every input. We also have the following folklore lemma, proved the same way as Lemma 8, which states that UFAs are simulated by unambiguous protocols.

▶ **Lemma 14.** *If a two-party function $F\colon \{0,1\}^{m_1} \times \{0,1\}^{m_2} \to \{0,1\}$ admits an UFA with s states, then $\mathrm{Par}_1(F) \leq s$ (that is, $\mathrm{U}_1(F) \leq \log s$).* ◀

Nonnegative rank. We often think of a two-party function $F\colon X \times Y \to \{0,1\}$ as a boolean matrix $F \in \{0,1\}^{X\times Y}$, sometimes called the *communication matrix* of F. For a nonnegative matrix $M \in \mathbb{R}^{X\times Y}_{\geq 0}$ we define its *nonnegative rank*, denoted $\mathrm{rk}^+(M)$, as the least r such that M can be written as a sum of r nonnegative rank-1 matrices, i.e., $M = \sum_{i=1}^r u_i v_i^{\mathrm{T}}$, where $u_i \in \mathbb{R}^X_{\geq 0}$ and $v_i \in \mathbb{R}^Y_{\geq 0}$ are nonnegative vectors. Note that for a boolean matrix F,

$$\mathrm{Par}_1(F) \;\geq\; \mathrm{rk}^+(F) \qquad \text{and thus} \qquad \mathrm{U}_1(F) \;\geq\; \log \mathrm{rk}^+(F). \tag{2}$$

Indeed, if $F^{-1}(1)$ can be partitioned into r rectangles, $F^{-1}(1) = R_1 \sqcup \cdots \sqcup R_r$, then F can be written as a sum of r nonnegative rank-1 matrices, $F = M_1 + \cdots + M_r$, where M_i is 1 on the rectangle R_i and 0 elsewhere. As with nonnegative degree, we define an approximate version of nonnegative rank. The *ϵ-approximate nonnegative rank* of M, denoted $\mathrm{rk}^+_\epsilon(M)$, is defined as the least $\mathrm{rk}^+(N)$ over all nonnegative matrices N that ϵ-approximate M, i.e.,

$$|M_{ij} - N_{ij}| \;\leq\; \epsilon \qquad \text{for all } i, j.$$

Nonnegative lifting. Finally, we formulate a theorem that lifts lower bounds on the nonnegative degree of an n-bit boolean function f to the nonnegative rank of the composed function $F = f \circ g^n$ (which was defined in Section 2.1).

▶ **Theorem 15** ([12, 17]). *Fix constants $\delta > \epsilon > 0$. For any $n \in \mathbb{N}$ there is a gadget $g\colon \{0,1\}^b \times \{0,1\}^b \to \{0,1\}$ with $b = \Theta(\log n)$ such that for any $f\colon \{0,1\}^n \to \{0,1\}$ we have*

$$\log \mathrm{rk}^+_\epsilon(f \circ g^n) \;\geq\; \Omega(\deg^+_\delta(f) \cdot b).$$ ◀

3.4 Proof of Theorem 2 (and also Theorem 4)

We start with the function $f\colon \{0,1\}^n \to \{0,1\}$ given by Theorem 9 such that for $m = \mathrm{poly}(n)$,

$$\mathrm{UC}_1(f) \;\leq\; m, \tag{3}$$
$$\deg^+_{1.5\times 10^{-5}}(f^\vee) \;\geq\; \tilde\Omega(m^2). \tag{4}$$

We then use the gadget g on $b = \Theta(\log n)$ bits from the lifting theorem Theorem 15 to construct $F := f \circ g^n$. By the same argument as in Section 2.2 we see that the resulting $F\colon \{0,1\}^{nb} \times \{0,1\}^{nb} \to \{0,1\}$ enjoys the following upper bounds, derived from (3).

- F admits an unambiguous DNF of width $2bm = \tilde O(m)$.
- F admits an UFA of size $2^{\tilde O(m)}$.
- F admits an unambiguous protocol of cost $\mathrm{U}_1(F) \leq \tilde O(m)$.

On the other hand, we note that $F^\vee = (f \circ g^n)^\vee = f^\vee \circ g^n$. Hence, we may combine (2), Theorem 15, and (4) to conclude that

$$\mathrm{U}_1(F^\vee) \;\geq\; \log \mathrm{rk}^+_{10^{-5}}(F^\vee) \;\geq\; \Omega(\deg^+_{1.5\times 10^{-5}}(f^\vee)) \;\geq\; \tilde\Omega(m^2). \tag{5}$$

This finishes the proof of Theorem 4. We proceed with the proof of Theorem 2. To this end, we define two languages

$$L_1 \;:=\; \{xx'yy' : x, x', y, y' \in \{0,1\}^{bn} \text{ and } F(x,y) = 1\},$$
$$L_2 \;:=\; \{xx'yy' : x, x', y, y' \in \{0,1\}^{bn} \text{ and } F(x',y') = 1\}.$$

Both L_1 and L_2 admit UFAs of size $\mathrm{poly}(n) \cdot 2^{\tilde O(m)} = 2^{\tilde O(m)} =: N$. By contrast, we have $L_1 \cup L_2 = (F^\vee)^{-1}(1)$, and this union language requires UFAs of size $2^{\tilde\Omega(m^2)} = N^{\tilde\Omega(\log N)}$ by (5) and Lemma 14. This concludes the proof of Theorem 2. ◀

4 UFA Separation

In this section, we prove Theorem 3.

▶ **Theorem 3** (Separation). *For every $n \in \mathbb{N}$ there is a language $L \subseteq \{0,1\}^*$ such that both L and $\overline{L}$ are recognised by n-state NFAs but any UFA that recognises L requires $n^{\Omega(\log n)}$ states.*

Loosely speaking, in our construction, we define NFAs $\mathcal{A}_1, \mathcal{A}_2$ that recognize *(sparse) set disjointness* and its complement. For $n \in \mathbb{N}$ and $k \leq n$ we define

$$\textsc{Disj}^n_k \;:=\; \{(S,T) \mid S \subseteq [n],\ T \subseteq [n],\ |S| = |T| = k,\ S \cap T = \emptyset\}\,.$$

Define also $\langle \textsc{Disj}^n_k \rangle := \{\langle S\rangle\langle T\rangle \mid (S,T) \in \textsc{Disj}^n_k\}$ where $\langle S\rangle \in \{0,1\}^n$ is such that the ith letter of $\langle S\rangle$ is 1 if and only if $i \in S$, and similarly for $\langle T\rangle$. Note that $\langle S\rangle, \langle T\rangle$ each contain k times the letter 1. To prove Theorem 3 it suffices to prove the following lemma.

▶ **Lemma 16.** *For any $n \in \mathbb{N}$ let $k := \lceil \log_2 n \rceil$. There are NFAs $\mathcal{A}_1, \mathcal{A}_2$ with $n^{O(1)}$ states such that $L(\mathcal{A}_1) = \langle \textsc{Disj}^n_k \rangle$ and $L(\mathcal{A}_2) = \{0,1\}^* \setminus \langle \textsc{Disj}^n_k \rangle$. Furthermore, any UFA that recognizes $\langle \textsc{Disj}^n_k \rangle$ has at least $n^{\Omega(\log n)}$ states.*

In the rest of the section we prove Lemma 16 by following Razborov's analysis of sparse set disjointness [21]. In particular, we will give a self-contained proof of the existence of polynomial-sized NFAs for $\langle \textsc{Disj}^n_k \rangle$ and its complement, but the main argument also comes from communication complexity.

4.1 Proof of Lemma 16

First we prove the statement on UFAs. Write $\binom{[n]}{k} := \{S \subseteq [n] \mid |S| = k\}$. Let $F \colon \binom{[n]}{k} \times \binom{[n]}{k} \to \{0,1\}$ be the two-party function with $F(S,T) = 1$ if and only if $(S,T) \in \textsc{Disj}^n_k$. It is shown, e.g., in [18, Example 2.12] that the communication matrix of F has full rank, $\mathrm{rk}(F) = \binom{n}{k}$. Let $F' \colon \{0,1\}^n \times \{0,1\}^n \to \{0,1\}$ be such that $F'(x,y) = 1$ if and only if $xy \in \langle \textsc{Disj}^n_k \rangle$. Then F is a principal submatrix of F', so $\binom{n}{k} \le \mathrm{rk}(F')$. Using (2) and Lemma 14 it follows that any UFA, say $\mathcal{A}$, that recognizes $\langle \textsc{Disj}^n_k \rangle$ has at least $\binom{n}{k} \ge (\frac{n}{k})^k$ states. With $k := \lceil \log_2 n \rceil$, it follows that $\mathcal{A}$ has $n^{\Omega(\log n)}$ states.

It is easy to see that there is an NFA, $\mathcal{A}_2$, with $n^{O(1)}$ states and $L(\mathcal{A}_2) = \{0,1\}^* \setminus \langle \textsc{Disj}^n_k \rangle$. Indeed, we can assume that the input is of the form $\langle S \rangle \langle T \rangle$; otherwise $\mathcal{A}_2$ accepts. NFA $\mathcal{A}_2$ guesses $i \in [n]$ such that $i \in S \cap T$ and then checks it.

Finally, we show that there is an NFA, $\mathcal{A}_1$, with $n^{O(1)}$ states and $L(\mathcal{A}_1) = \langle \textsc{Disj}^n_k \rangle$. We can assume that the input is of the form $\langle S \rangle \langle T \rangle$; otherwise $\mathcal{A}_1$ rejects. NFA $\mathcal{A}_1$ "hard-codes" polynomially many sets $Z_1, \ldots, Z_\ell \subseteq [n]$. It guesses $i \in [\ell]$ such that $S \subseteq Z_i$ and $Z_i \cap T = \emptyset$ and then checks it. It remains to show that there exist $\ell = n^{O(1)}$ sets $Z_1, \ldots, Z_\ell \subseteq [n]$ such that for any $(S,T) \in \textsc{Disj}^n_k$ there is $i \in [\ell]$ with $S \subseteq Z_i$ and $Z_i \cap T = \emptyset$. The argument uses the probabilistic method and is due to [21]; see also [18, Example 2.12]. We reproduce it here due to its elegance and brevity.

Fix $(S,T) \in \textsc{Disj}^n_k$. Say that a set $Z \subseteq [n]$ *separates* (S,T) if $S \subseteq Z$ and $Z \cap T = \emptyset$. A random set $Z \subseteq [n]$ (each i is in Z with probability $1/2$) separates (S,T) with probability 2^{-2k}. Thus, choosing $\ell := \lceil 2^{2k} \ln \binom{n}{k}^2 \rceil = n^{O(1)}$ random sets $Z \subseteq [n]$ independently, the probability that none of them separates (S,T) is

$$(1 - 2^{-2k})^\ell \;<\; e^{-2^{-2k}\ell} \;\le\; \binom{n}{k}^{-2}.$$

By the union bound, since $|\textsc{Disj}^n_k| < \binom{n}{k}^2$, the probability that there exists $(S,T) \in \textsc{Disj}^n_k$ such that none of ℓ random sets separates (S,T) is less than 1. Equivalently, the probability that for all $(S,T) \in \textsc{Disj}^n_k$ at least one of ℓ random sets separates (S,T) is positive. It follows that there are $Z_1, \ldots, Z_\ell \subseteq [n]$ such that each $(S,T) \in \textsc{Disj}^n_k$ is separated by some Z_i. ◀

5 Bonus result: Approximate nonnegative rank

In this section, we prove Theorem 5.

▶ **Theorem 5** (No efficient error reduction). *For every $m \in \mathbb{N}$ there exists a boolean matrix M with $\mathrm{rk}^+_{1/4}(M) \le m$ but such that $\mathrm{rk}^+_{10^{-5}}(M) \ge m^{\tilde{\Omega}(\log m)}$.*

We first illustrate the idea in the context of nonnegative degree. In contrast to Theorem 9 (which states that $\vee$ is hard to approximate to within tiny error), we show that the $\vee$-operation is, in fact, easy to approximate when we allow large enough error.

▷ Claim 17. For any boolean-valued f, we have $\deg^+_{1/4}(f^\vee) \le \deg^+(f)$.

Proof. Let $g \colon \{0,1\}^{2n} \to \mathbb{R}_{\ge 0}$ be given by $g(x,y) := (f(x) + f(y))/2 + 1/4$. Then

$$g(x,y) \;=\; \begin{cases} 1/4 & \text{if } f(x) = f(y) = 0, \\ 5/4 & \text{if } f(x) = f(y) = 1, \\ 3/4 & \text{otherwise.} \end{cases}$$

Thus g is a $1/4$-approximation to $f^\vee$. Note also that $\deg^+(g) \le \deg^+(f)$, as desired. ◁

We can now repeat the same idea for nonnegative rank. In Section 3.4 we constructed a boolean matrix (two-party function) F such that $\log \mathrm{rk}^+(F) \leq \mathrm{U}_1(F) \leq m$ and $\log \mathrm{rk}^+_{10^{-5}}(F^\vee) \geq \tilde{\Omega}(m^2)$. We claim that $\log \mathrm{rk}^+_{1/4}(F^\vee) \leq O(m)$, which would finish the proof of Theorem 5. Indeed, analogously to Claim 17, we can define a nonnegative matrix by $G(xx, yy') \coloneqq (F(x,y) + F(x',y'))/2 + 1/4$. This is a 1/4-approximation to $F^\vee$ and we have $\mathrm{rk}^+(G) \leq 2 \cdot \mathrm{rk}^+(F) + 1 \leq 2^{m+1} + 1$, as claimed.

References

1 Noga Alon. Problems and results in extremal combinatorics–I. *Discrete Mathematics*, 273(1–3):31–53, 2003. doi:10.1016/S0012-365X(03)00227-9.

2 Kaspars Balodis, Shalev Ben-David, Mika Göös, Siddhartha Jain, and Robin Kothari. Unambiguous DNFs and Alon-Saks-Seymour. In *62nd IEEE Annual Symposium on Foundations of Computer Science (FOCS)*, 2021. To appear. Available at arXiv:2102.08348.

3 Jean-Camille Birget. Partial orders on words, minimal elements of regular languages, and state complexity. *Theoretical Computer Science*, 119(2):267–291, 1993. doi:10.1016/0304-3975(93)90160-U.

4 Mikołaj Bojańczyk and Wojciech Czerwiński. An automata toolbox, 2018. Available at https://www.mimuw.edu.pl/~bojan/paper/automata-toolbox-book.

5 Thomas Colcombet. Unambiguity in automata theory. In *17th International Workshop on Descriptional Complexity of Formal Systems (DCFS)*, volume 9118 of *Lecture Notes in Computer Science*, pages 3–18. Springer, 2015. doi:10.1007/978-3-319-19225-3_1.

6 Thomas Colcombet, Karin Quaas, and Michał Skrzypczak. Unambiguity in Automata Theory (Dagstuhl Seminar 21452). *Dagstuhl Reports*, 11(10):57–71, 2022. doi:10.4230/DagRep.11.10.57.

7 Wojciech Czerwiński, Laure Daviaud, Nathanaël Fijalkow, Marcin Jurdziński, Ranko Lazić, and Pawel Parys. Universal trees grow inside separating automata: Quasi-polynomial lower bounds for parity games. In *Proceedings of the 2019 Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 2333–2349, 2019. doi:10.1137/1.9781611975482.142.

8 Wojciech Czerwiński and Sławomir Lasota. Regular separability of one counter automata. *Logical Methods in Computer Science*, 15(2), 2019. doi:10.23638/LMCS-15(2:20)2019.

9 Yuan Gao, Nelma Moreira, Rogério Reis, and Sheng Yu. A survey on operational state complexity. *Journal of Automata, Languages and Combinatorics*, 21(4):251–310, 2016. doi:10.25596/jalc-2016-251.

10 Mika Göös. Lower bounds for clique vs. independent set. In *IEEE 56th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 1066–1076. IEEE Computer Society, 2015. doi:10.1109/FOCS.2015.69.

11 Mika Göös, T.S. Jayram, Toniann Pitassi, and Thomas Watson. Randomized communication versus partition number. *ACM Transactions on Computation Theory (TOCT)*, 10(1):1–20, 2018. doi:10.1145/3170711.

12 Mika Göös, Shachar Lovett, Raghu Meka, Thomas Watson, and David Zuckerman. Rectangles are nonnegative juntas. *SIAM Journal on Computing*, 45(5):1835–1869, 2016. doi:10.1137/15M103145X.

13 Emil Indzhev and Stefan Kiefer. On complementing unambiguous automata and graphs with many cliques and cocliques. Technical report, arxiv.org, 2021. Available at arXiv:2105.07470.

14 Galina Jirásková. State complexity of some operations on binary regular languages. *Theoretical Computer Science*, 330(2):287–298, 2005. doi:10.1016/j.tcs.2004.04.011.

15 Jozef Jirásek Jr., Galina Jirásková, and Juraj Šebej. Operations on unambiguous finite automata. *International Journal of Foundations of Computer Science*, 29(5):861–876, 2018. doi:10.1142/S012905411842008X.

16 Gillat Kol, Shay Moran, Amir Shpilka, and Amir Yehudayoff. Approximate nonnegative rank is equivalent to the smooth rectangle bound. *Computational Complexity*, 28(1):1–25, 2019. Preliminary version in *ICALP '14*. doi:10.1007/s00037-018-0176-4.

17 Pravesh Kothari, Raghu Meka, and Prasad Raghavendra. Approximating rectangles by juntas and weakly exponential lower bounds for LP relaxations of CSPs. *SIAM Journal on Computing*, pages STOC17–305–STOC17–332, May 2021. Preliminary version in *STOC '17*. doi:10.1137/17m1152966.

18 Eyal Kushilevitz and Noam Nisan. *Communication Complexity*. Cambridge University Press, 1997. doi:10.1017/CBO9780511574948.

19 Anup Rao and Amir Yehudayoff. *Communication Complexity: And Applications*. Cambridge University Press, 2020. doi:10.1017/9781108671644.

20 Mikhail Raskin. A superpolynomial lower bound for the size of non-deterministic complement of an unambiguous automaton. In *45th International Colloquium on Automata, Languages, and Programming (ICALP 2018)*, volume 107 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 138:1–138:11, 2018. doi:10.4230/LIPIcs.ICALP.2018.138.

21 Alexander Razborov. Applications of matrix methods to the theory of lower bounds in computational complexity. *Combinatorica*, 10(1):81–93, 1990. doi:10.1007/BF02122698.

22 Wikipedia. State complexity. URL: https://en.wikipedia.org/wiki/State_complexity.

23 Mihalis Yannakakis. Expressing combinatorial optimization problems by linear programs. *Journal of Computer and System Sciences*, 43(3):441–466, 1991. doi:10.1016/0022-0000(91)90024-Y.