Attainable Values of Reset Thresholds

Michalina Dżyga[1], Robert Ferens[*2], Vladimir V. Gusev[†3], and Marek Szykuła[‡4]

1 **Institute of Computer Science, University of Wrocław, Joliot-Curie 15, Wrocław, Poland**
misia.sieradzka@interia.pl

2 **Institute of Computer Science, University of Wrocław, Joliot-Curie 15, Wrocław, Poland**
robert.ferens@interia.pl

3 **ICTEAM Institute, Université catholique de Louvain, Louvain-la-Neuve, Belgium, and Institute of Natural Sciences and Mathematics, Ural Federal University, Ekaterinburg, Russia**
vl.gusev@gmail.com

4 **Institute of Computer Science, University of Wrocław, Joliot-Curie 15, Wrocław, Poland**
msz@cs.uni.wroc.pl

Abstract

An automaton is synchronizing if there exists a word that sends all states of the automaton to a single state. The reset threshold is the length of the shortest such word. We study the set RT_n of attainable reset thresholds by automata with n states. Relying on constructions of digraphs with known local exponents we show that the intervals $\left[1,(n^2-3n+4)/2\right]$ and $[(p-1)(q-1), p(q-2)+n-q+1]$, where $2 \le p < q \le n, p+q > n, \gcd(p,q) = 1$, belong to RT_n, even if restrict our attention to strongly connected automata. Moreover, we prove that in this case the smallest value that does not belong to RT_n is at least $n^2 - O(n^{1.7625} \log n / \log\log n)$. This value is increased further assuming certain conjectures about the gaps between consecutive prime numbers. We also show that any value smaller than $\frac{n(n-1)}{2}$ is attainable by an automaton with a sink state and any value smaller than $n^2 - O(n^{1.5})$ is attainable in general case. Furthermore, we solve the problem of existence of slowly synchronizing automata over an arbitrarily large alphabet, by presenting for every fixed size of the alphabet an infinite series of irreducibly synchronizing automata with the reset threshold $n^2 - O(n)$.

1998 ACM Subject Classification F.1.1 Models of Computation, G.2.2 Graph Theory

Keywords and phrases Černý conjecture, exponent, primitive digraph, reset word, synchronizing automaton

Digital Object Identifier 10.4230/LIPIcs.MFCS.2017.40

* R. Ferens was supported in part by the National Science Centre, Poland under project number 2014/15/B/ST6/00615.
† V. Gusev was supported by the French Community of Belgium, by the IAP network DYSCO, RFBR grant no. 16-01-00795, Russian Ministry of Education and Science project no. 1.3253.2017, and the Competitiveness Enhancement Program of Ural Federal University.
‡ M. Szykuła was supported in part by the National Science Centre, Poland under project number 2014/15/B/ST6/00615.

1 Introduction

Let $\mathscr{A} = (Q, \Sigma, \delta)$ be a *deterministic complete finite (semi)automaton* with the set of states Q, the alphabet Σ, and the transition function $\delta\colon Q \times \Sigma \to Q$. We denote the image of a state q under the action of a word w as $q \cdot w$. Automaton $\mathscr{A}$ is called *synchronizing* if there exist a word w and a state f such that for every state $q \in Q$ we have $q \cdot w = f$. Every such word w is called *synchronizing* (or *reset*) word for $\mathscr{A}$. The length of the shortest synchronizing word of $\mathscr{A}$ is called the *reset threshold* and is denoted by $\mathrm{rt}(\mathscr{A})$. This notion can be extended to subsets: for a subset $S \subseteq Q$, $\mathrm{rt}(\mathscr{A}, S)$ is the length of the shortest word w such that $|\delta(S, w)| = 1$.

Synchronizing automata constitute a well-studied class of automata with applications to group theory [4], coding theory [6, Chapter 10], industrial automation, matrix and control theories [7], etc. A brief survey of the theory of synchronizing automata can be found in [18, 33].

The reset threshold is one of the most important characteristics of a synchronizing automaton. One can compute a synchronizing word following a greedy strategy in polynomial-time, but finding the shortest such word is hard. The problem of deciding whether $\mathrm{rt}(\mathscr{A}) = k$ for a given automaton $\mathscr{A}$ and an integer k encoded in binary is DP-complete [23]. Moreover, unless P = NP, there is no polynomial-time algorithm computing the reset threshold with the approximation error $O(n^{1-\varepsilon})$ for a fixed $\varepsilon > 0$ [13]. Furthermore, widely used greedy algorithms have approximation error $\Omega(n)$, where n is the number of states in a given automaton [2].

One of the most famous open problems in automata theory is to determine general bounds on the reset thresholds of automata with n states. *The Černý conjecture* states that the reset threshold of an automaton is at most $(n-1)^2$ [10, 11]. Furthermore, this bound is reached by the Černý automaton $\mathscr{C}_n$ with n states, see [33, p. 18]. Despite intensive efforts of researchers and confirmation of this conjecture in various special classes of automata the best upper bound obtained so far is $(15617n^3 + 7500n^2 + 9375n - 31250)/93750$ [29].

The difficulty of the Černý conjecture led to a large number of attempts to disprove it by means of a counterexample obtained via computational experiments [31, 3, 20]. Although, no counterexample was ever found, several interesting observations were made. For example, it was noted that there are no synchronizing automata with n states and two letters with the reset threshold in the range $[n^2 - 3n + 5, (n-1)^2 - 1]$ for $n = 7, \ldots, 12$. Afterwards, a few more potential "gaps" were identified. It leads us in turn to the following general question that we address in our paper:

> What is the set RT_n of possible reset thresholds
> for synchronizing automata with n states?

Clearly, a complete answer to this question is out of reach for the state of the art techniques as the Černý conjecture is merely about an upper bound on RT_n. Nevertheless, it is possible to describe certain values belonging to RT_n by presenting synchronizing automata with known reset thresholds. This problem is also not trivial as one has to prove that a constructed automaton is not only synchronized by a word of the claimed length, but also demonstrate that any shorter word is not synchronizing.

Additional interest to results of this kind comes from the following reasons. Concrete examples of synchronizing automata with known reset thresholds shed light on the phenomenon of synchronization giving hints on potential proof of the Černý conjecture. Furthermore, the availability of hands on examples is important for the evaluation of new ideas and algorithms.

Synchronizing automata with the reset threshold close to $(n-1)^2$, so called slowly synchronizing automata, are especially important as they almost never appear in random samples of synchronizing automata. Moreover, the ideas used in construction of synchronizing automata with known reset thresholds could also lead to interesting connections between different fields. For example, a new line of research devoted to the interplay between synchronizing automata and primitive matrices was started in [3], see [14] for recent results. Part of our current work can be seen as a continuation of this line of research.

Over the years a large number of synchronizing automata with known reset thresholds were presented in the literature. It can be either sporadic examples possessing interesting properties [31, 33] or infinite series of automata designed to give a lower bound on the largest reset threshold among synchronizing automata belonging to a special class. For example, constructions of automata with the sink state are given in [1, 24, 26], Eulerian automata in [15, 30], automata with the reset threshold close to $(n-1)^2$ [3]. Multi-parametric series of automata with the aim of covering RT_n were presented in [16].

Our contributions. In the present paper we significantly expand the list of values known to belong to RT_n. Moreover, we present non-trivial lower bounds on the smallest value that does not belong to RT_n. Our approach can be summarized as follows. Recall that a subset of states $\mathcal{C}$ of the automaton $\mathscr{A}$ is a *sink* if $\delta(q,\ell) \in \mathcal{C}$ for every $q \in \mathcal{C}$ and $\ell \in \Sigma$. A classical observation states that a synchronizing automaton has the unique strongly connected sink component $\mathcal{C}$. Furthermore, the process of synchronization can be performed in the following manner: initially, one applies a word u such that $Q \cdot u \subseteq \mathcal{C}$, i.e. the automaton is brought to $\mathcal{C}$, afterwards, one applies v such that $|\mathcal{C} \cdot v| = 1$. Therefore, we consider three natural cases.

Case (i) strongly connected automata. In Section 2 we construct synchronizing automata belonging to this class with known reset thresholds based on examples of digraphs with known local exponents [28]. We show that the intervals $\left[1, (n^2-3n+4)/2\right]$ and $[(p-1)(q-1), p(q-2)+n-q+1]$, where $2 \le p < q \le n, p+q > n, \gcd(p,q) = 1$, belong to the set of attainable reset thresholds by n-state automata. The method that we use to convert digraphs into synchronizing automata is based on techniques recently introduced in [7, 14].

Let $\mathrm{gt}_{sc}(n)$ be the smallest value that does not serve as the reset threshold of a strongly connected automaton with n states. In Section 3 we show that $\mathrm{gt}_{sc}(n)$ is at least $n^2 - \tilde{O}(n^{1.7625})$, where $\tilde{O}(f(n))$ is the shorthand for $O(f(n)\log^k(f(n))$ for some k. Moreover, we strengthen this bound conditioning on the validity of classical conjectures related to the distribution of prime numbers. If the Riemann hypothesis is true, then $\mathrm{gt}_{sc}(n) \ge n^2 - \tilde{O}(n^{1.75})$. If the Cramer's conjecture is true, then $\mathrm{gt}_{sc}(n) \ge n^2 - \tilde{O}(n^{1.5})$. Our proofs are based on rigorous analysis of the overlaps of the aforementioned intervals. Similar intervals often appear in index set problems about digraphs and Boolean matrices, so our techniques can be applied to them as well [22].

Case (ii) automata with the sink state. The reset threshold of automata belonging to this class is bounded by $\frac{n(n-1)}{2}$, moreover, there exists a series of n-state automata reaching this bound [26]. In Section 4 we generalize this series to prove that for every $1 \le \ell \le \frac{n(n-1)}{2}$ there exists an n-state automaton with sink and the reset threshold equal to ℓ.

Case (iii) automata without restrictions. In Section 5 we utilize ideas of the previous cases to show that for every $1 \le \ell \le n^2 - O(n^{1.5})$ there exists an n-state synchronizing automaton with reset threshold equal to ℓ. This result does not depend on the number theoretic conjectures.

The number of letters of automata constructed in the previous cases grows with the number of states. In Section 6 we aim to get a better understanding on how the number of letters influences the set of possible reset thresholds. To avoid trivial cases we focus on *irreducibly synchronizing* automata, i.e. automata that become non-synchronizing after the removal of any letter. We resolve an open problem asking whether for each fixed size of the alphabet there is a series of irreducibly synchronizing n-state automata with the reset threshold $n^2 - O(n)$. Previously, such automata were known only over 2-letter and 3-letter alphabets [3, 21]. Namely, we construct infinite (in n and k) series of automata $\mathscr{M}_{n,k}$, $\mathscr{M}'_{n,k}$ with n states and k letters such that $rt(\mathscr{M}_{n,k}) = n^2 - (k+3)n + 2k + 3$ and $rt(\mathscr{M}'_{n,k}) = n^2 - (k+3)n + 2k + 4$. These examples can be also seen as a formal bound to the following common empirical statement: synchronizing automata with large number of letters have relatively small reset thresholds (due to a large number of possibilities at every step of synchronization).

2 Strongly connected automata

A digraph G is *primitive* if there exists a positive integer t such that for every pair of vertices u, v of G there exists a walk from u to v of length exactly t. The smallest such t is called the exponent of G and denoted by $\exp(G)$. A survey of results about this classical notion can be found in [9, Chapter 3.5].

The notion of local exponent was introduced in [8]. The *local exponent* of G at a vertex u, denoted by $\exp_G(u)$ or $\exp(u)$, is the smallest t such that for every vertex v of G there is a walk from u to v of length exactly t. Let $V = \{1, 2, \ldots, n\}$. We will always assume that the vertices are reordered so that $\exp_G(1) \le \exp_G(2) \le \ldots \le \exp_G(n)$.

The behavior of the exponents and the local exponents of digraphs with n vertices gained a lot of attention in literature. Let $\mathrm{ES}_n(1)$ be the set of possible first local exponents of all digraphs with n vertices, i.e. $\mathrm{ES}_n(1) = \{\exp_G(1) \mid G = (V, E)\,, |V| = n\,, G \text{ is primitive}\}$.

▶ **Theorem 1** ([28, Theorem 9]).

$$\mathrm{ES}_n(1) = \left[1, \frac{n^2 - 3n + 4}{2}\right] \quad \bigcup \quad \bigcup_{(p,q)\in L(n)} [(p-1)(q-1), p(q-2) + n - q + 1],$$

where $L(n) = \{(p, q) : 2 \le p < q \le n, p + q > n, \gcd(p, q) = 1\}$.

We will rely on Theorem 1 to construct synchronizing automata with known reset thresholds. The proof of the following proposition is based on the "determinization" procedure appearing in [7, 14, 17].

▶ **Proposition 2.** *Let $G(V, E)$ be a primitive n-vertex digraph. Then there exists a synchronizing n-state automaton $\mathscr{A}$ such that* $\mathrm{rt}(\mathscr{A}) = \exp_G(1)$.

Proof. The automaton $\mathscr{A}$ is constructed as follows. The set of states is equal to $V = \{1, 2, \ldots, n\}$. For every choice of states $s_1, s_2, \ldots, s_n \in V$ such that $(s_1, 1), (s_2, 2), \ldots, (s_n, n) \in E$ we add the letter $(s_1, \ldots, s_n)$ with the action $\delta(j, (s_1, \ldots, s_n)) = s_j$ for every $j \in V$.

We need to show that $\mathscr{A}$ is synchronizing and $\mathrm{rt}(\mathscr{A}) = \exp_G(1)$. Since G is primitive and every vertex is reachable from 1 in $\exp_G(1)$ steps there exists a sequence of n-tuples $(v_1^{(1)} = 1, v_2^{(1)} = 1, \ldots, v_n^{(1)} = 1), (v_1^{(2)}, v_2^{(2)}, \ldots, v_n^{(2)}), \ldots, (v_1^{(t)} = 1, v_2^{(t)} = 2, \ldots, v_n^{(t)} = n)$ of length $t = \exp_G(1)$ and such that for every $2 \le i \le t$, $1 \le j \le n$ we have $(v_j^{(i-1)}, v_j^{(i)}) \in E$. Furthermore, we can assume that if $v_j^{(i)} = v_k^{(i)}$ for some i, j, k then $v_j^{(\ell)} = v_k^{(\ell)}$ for all $\ell \le i$.

Indeed, by substituting the value of $v_j^{(\ell)}$ to $v_k^{(\ell)}$ for all $\ell \leq i$ we will obtain a sequence satisfying all of the aforementioned properties.

Observe now that for every $i \geq 2$ there is a letter of $\mathscr{A}$ that maps the tuple $(v_1^{(i)}, v_2^{(i)}, \ldots, v_n^{(i)})$ to $(v_1^{(i-1)}, v_2^{(i-1)}, \ldots, v_n^{(i-1)})$, namely, any letter $(s_1, \ldots, s_n)$ satisfying $s_{v_j^{(i)}} = v_j^{(i-1)}$ for all $1 \leq j \leq n$. Thus, the sequence of tuples can be seen as an application of a word w mapping the set $\{1, 2, \ldots n\}$ to $\{1\}$. In other words, w is a synchronizing word of length $\exp_G(1)$.

It remains to note that $\mathrm{rt}(\mathscr{A}) \geq \exp_G(1)$. Indeed, every synchronizing word w mapping V to $\{f\}$ labels walks leading from every state to f; moreover, the edges of $\mathscr{A}$ are the inverted edges of G. Thus, every vertex of G is reachable from f in $|w|$ steps. Since $\exp_G(1)$ is the smallest number with this property, we have $|w| \geq \exp_G(1)$. ◀

By combining Theorem 1 and Proposition 2 we obtain the main result of this section:

▶ **Theorem 3.** *For every n, $\mathrm{ES}_n(1) \subset \mathrm{RT}_n$. Furthermore, it remains true even in the case of strongly connected automata.*

▶ **Remark 4.** *Clearly, $\mathrm{RT}_n \not\subset \mathrm{ES}_n(1)$, since the largest local 1-exponent is at most $n^2 - 3n + 3$, by Theorem 1 (originally [27, Theorem 2.1]), while the Černý series of automata has the reset threshold equal to $(n-1)^2$.*

Proposition 2 is also tightly connected to the Hybrid Černý-Road Coloring conjecture [3, Conjecture 2]. Let G be a digraph with the set of edges E, and Σ be a finite alphabet. A *coloring* of G is an arbitrary deterministic finite state automaton obtained by distributing letters of Σ over the edges E. Note that G typically has a large number of colorings. The celebrated Road Coloring Theorem states that every primitive digraph with out-degree k has a synchronizing coloring with k letters [32]. The Hybrid Černý-Road Coloring conjecture states that such synchronizing coloring can always be found with the reset threshold at most $n^2 - 3n + 3$.

▶ **Corollary 5.** *Let G be a primitive digraph with n vertices. There exists an alphabet Σ and a coloring $\mathscr{A}$ of G with Σ such that $\mathscr{A}$ is a synchronizing automaton and $\mathrm{rt}(\mathscr{A}) \leq n^2 - 3n + 3$. In other words, the Hybrid Černý-Road Coloring conjecture holds true if we are allowed to use an alphabet of arbitrary size.*

Proof. The proof of Proposition 2 describes a procedure to derive a synchronizing coloring of $\overleftarrow{G}$ – the digraph obtained by reversing all the edges of G, with the reset threshold equal to $\exp_G(1)$. Since $\exp_G(1)$ is bounded by $n^2 - 3n + 3$ by Theorem 1, the corollary follows. ◀

3 Lower bounds on the smallest unattainable value

In this section we will derive a lower bound on the smallest value that is not a reset threshold of a strongly connected n-state automaton. The proof of main theorem is based on the following number theoretic result.

Let $g(x)$ be the maximal difference (*prime gap*) between any prime number $p \leq x$ and the next prime number. It is known that $g(x) \leq x^{0.525}$ when x is large enough [5].

Let $\omega(x)$ be the maximal number of distinct prime divisors of any number $i \leq x$. It is known that $\omega(x) \leq 1.38402 \log x / \log\log x$ for $x \geq 3$ [25].

▶ **Theorem 6.** *For n large enough, the function*

$$f(n) = 6n + 4n \cdot g(3\sqrt{n}) \cdot \omega(n) \cdot \left(g(3\sqrt{n}) \cdot \omega(n) + 6\sqrt{n}\right)$$

satisfies the following equation:

$$\left[\left\lceil\frac{n^2}{3}\right\rceil, \lfloor n^2 - f(n)\rfloor\right] \subseteq \bigcup_{(p,q)\in L(n)} [(p-1)(q-1), (p-1)(q-1)+n-p],$$

where $L(n) = \{(p,q) \mid 2 \le p < q \le n, p+q > n, \gcd(p,q) = 1\}$.

▶ **Corollary 7.** *Using the known upper bound* $g(x) \le x^{0.525}$ [5] *we obtain*

$$f(n) = O(n^{1.7625} \log n / \log\log n).$$

Moreover, if we assume the Riemann hypothesis, which implies $g(x) \in O(\sqrt{x}\log x)$ [12], *we get*

$$f(n) = O(n^{1.75} \log^2 n / \log\log n).$$

If we assume the Cramer's conjecture $g(x) \in O(\log^2 x)$ [12], *we get*

$$f(n) = O(n^{1.5} \log^3 n / \log\log n).$$

Before proving Theorem 6 we will derive the main result of this section:

▶ **Theorem 8.** *Let* $\mathrm{gt}_{sc}(n) \ge 1$ *be the smallest unattainable value by the reset thresholds of* n*-state strongly connected synchronizing automata. Then* $\mathrm{gt}_{sc}(n)$ *grows at least as fast as* $n^2 - O(f(n))$, *where* $f(n)$ *is the function from Corollary 7. In particular,*

$$\lim_{n\to\infty} \mathrm{gt}_{sc}(n)/n^2 = 1.$$

Proof. By Theorems 1 and 3 and Proposition 2 we know that for every integer in $[1, (n^2-3n+4)/2]$ there exists an automaton with the reset threshold equal to it. Also, from Theorem 6 and Corollary 7 we obtain the same result for the interval $[\left\lceil\frac{n^2}{3}\right\rceil, n^2 - O(n^{1.7625}\log n/\log\log n)]$. In other words, $\mathrm{gt}_{sc}(n)$ is at least $n^2 - O(f(n))$. Since $O(n^{1.7625}\log n/\log\log n)$ grows strictly slower than n^2, we have $\mathrm{gt}_{sc}(n)/n^2$ tending to 1 when $n \to \infty$. ◀

▶ **Remark 9.** *Theorem 6 establishes a lower bound on the least value that does not belong to* ES_n *as well.*

Now we are going to prove Theorem 6. Let n be large enough. Let x be any integer from $[\lceil\frac{n^2}{3}\rceil, \lfloor n^2 - f(n)\rfloor]$. We will show that x falls in the interval $[(p-1)(q-1), (p-1)(q-1)+n-p]$ of some $(p,q) \in L(n)$. Let k be the smallest odd integer such that $x \le k^2$.

$$(k-2)^2 < x \le k^2.$$

We define

$$S = \{p \in [3, n-k-1] \mid (p \text{ is prime}) \wedge p \nmid (k+1)\}.$$

▶ **Lemma 10.** *For every* $s \in S$, *each pair* $(p,q) = (k+1-s, k+1+s)$ *is in* $L(n)$.

Proof. It is enough to check all conditions for (p,q) in $L(n)$. For the first condition $p = k+1-s \ge k+1-(n-k-1) = 2(k+1)-n \ge 2\sqrt{x}-n$. Because $3x \ge n^2$ (so $\sqrt{x} \ge n/\sqrt{3}$), we have $2\sqrt{x} - n \ge (2/\sqrt{3}-1)n$. So for $n \ge 13$, we have $p \ge 2$.

The condition $p < q$ is obvious, and $q = k+1+s \le k+1+n-k-1 = n$.

The next condition states that $p+q>n$ and it is clear, since $p+q = 2(k+1) \geq 2\sqrt{x} \geq 2n/\sqrt{3} > n$.

Now we show that p and q are coprime. Let d be a non-trivial common divisor of p and q. Then also d divides $q-p=2s$. Since s is prime, d is either 2 or s. But $q=k+1+s$ is odd, since $k+1$ is even and s is odd. Also s cannot divide $q=k+1+s$, since by definition of s, s is not a divisor of $k+1$. Thus, p and q are coprime. ◀

Let $\mathrm{next}_S(i)$ be the smallest number in S greater than i (if it exists). Let $\mathrm{next}_P(i)$ be similarly defined for the set of prime numbers. For all i, we have $\mathrm{next}_P(i) - i \leq g(i)$.

To simplify formulas we define:

$$\begin{aligned}
b_k(s) &= (k-s)(k+s) = k^2 - s^2 \\
e_k(s) &= (k-s)(k+s) + n - (k+1) + s = k^2 - s^2 + n - (k+1) + s \\
I_k(s) &= [b_k(s), e_k(s)] = [k^2 - s^2, k^2 - s^2 + n - (k+1) + s]
\end{aligned}$$

Notice that $I_k(s)$ is the interval from Theorem 6 with $(p,q) = (k+1-s, k+1+s)$, which according to Lemma 10 is a pair from $L(n)$ for every $s \in S$.

Our plan is as follows. We will show the existence of $s_{max} \in S$ such that $b_k(s_{max}) \leq (k-2)^2$ and for every two consecutive elements $s, \mathrm{next}_S(s) \in S \cap [3, s_{max}]$, the intervals $I_k(s)$ and $I_k(\mathrm{next}_S(s))$ overlap. Since $b_k(s_{max}) \leq (k-2)^2$ and $k^2 \leq e_k(\mathrm{next}_S(0))$, it will prove that the intervals

$$I_k(\mathrm{next}_S(0)), \ldots, I_k(s), I_k(\mathrm{next}_S(s)), \ldots, I_k(s_{max})$$

cover all integers from $[(k-2)^2, k^2]$. So in particular x will be covered.

▶ **Lemma 11.** *For every $\varepsilon \in [0,1)$, when n is large enough, there is $s_{max} \in S$ satisfying $b_k(s_{max}) \leq (k-2)^2$ and $2\sqrt{k-1} \leq s_{max} \leq (2+\varepsilon)\sqrt{k-1}$.*

Proof. Let s_{max} be the smallest prime number $\geq 2\sqrt{k-1}$ and such that $s_{max} \nmid k+1$.

We show that $s_{max} \in S$. Let p be the first prime number $\geq 2\sqrt{k-1}$. If $p \neq s_{max}$, then $p \mid k+1$. But since $p \geq 2\sqrt{k-1}$ there is no other prime number $p' > p$ such that $p' \mid (k+1)$, as otherwise $p' \cdot p \geq 4(k-1) > k+1$ (for $k \geq 2$). Hence, $s_{max} = \mathrm{next}_P(p) > p$. So for both cases we get the upper bound:

$$\begin{aligned}
s_{max} &\leq \mathrm{next}_P(\mathrm{next}_P(2\sqrt{k-1})) \\
&\leq 2\sqrt{k-1} + 2g(2\sqrt{k-1} + g(2\sqrt{k-1})).
\end{aligned}$$

According to the bound $g(n) < n^{\theta}$ for $\theta = 0.525$ [5],

$$\begin{aligned}
s_{max} &\leq 2\sqrt{k-1} + 2(4\sqrt{k-1})^{\theta} \\
&\leq 2\sqrt{k-1} + 2 \cdot 4^{\theta} \cdot (k-1)^{\theta/2} \\
&\leq 2\sqrt{k-1}(1 + 4^{\theta} \cdot (k-1)^{(\theta-1)/2}).
\end{aligned}$$

Since $k \geq n/\sqrt{3}$, we have

$$0 \leq \lim_{k\to\infty} 4^{\theta}(k-1)^{0.5(\theta-1)} \leq \lim_{n\to\infty} 4^{\theta}(n/\sqrt{3}-1)^{0.5(\theta-1)} = 0.$$

So for n large enough, we obtain the upper bound:

$$s_{max} \leq (2+\varepsilon)\sqrt{k-1}.$$

Observe that:

$$\begin{aligned}
k^2 &= (k-2)^2 + 4k - 4 \\
&< x + 6k \\
&< n^2 - f(n) + 6n \\
&< n^2 - 24n\sqrt{n} \\
&< n^2 - 24n\sqrt{n} + (12\sqrt{n})^2,
\end{aligned}$$

and so $k < n - 12\sqrt{n}$. From this we obtain:

$$\begin{aligned}
s_{max} &\leq (2+\varepsilon)\sqrt{k-1} \\
&< 3\sqrt{n-1} \\
&< 12\sqrt{n} - 1 \\
&\leq n - (n - 12\sqrt{n}) - 1 \\
&\leq n - k - 1.
\end{aligned}$$

Thus, $s_{max} \in [3, n-k-1]$ and therefore is in S.
Finally, we have:

$$\begin{aligned}
s_{max} &\geq 2\sqrt{k-1} \\
s_{max}^2 &\geq 4(k-1) \\
k^2 - s_{max}^2 &\leq k^2 - 4k + 4 \\
b_k(s_{max}) &\leq (k-2)^2.
\end{aligned}$$

◀

▶ **Lemma 12.** *For every $s \in S \cap [3, s_{max} - 1]$, the intervals $I_k(s)$ and $I_k(\mathrm{next}_S(s))$ overlap, that is, $b_k(s) \leq e_k(\mathrm{next}_S(s))$. Moreover $k^2 \leq e_k(\mathrm{next}_S(0))$.*

Proof. We have $\mathrm{next}_S(s) \leq s_{max}$, because $s_{max} \in S$. Notice that

$$\mathrm{next}_S(s) \leq s + g(s_{max}) \cdot \big(\omega(k+1) + 1\big), \tag{1}$$

because the number of distinct prime odd divisors of $k+1$ is at most $\omega(k+1)$, and the gap between every two consecutive of the prime divisors is bounded from above by $g(s_{max})$.

Observe that:

$$\begin{aligned}
(k+1)^2 &< (k-2)^2 + 6k - 3 \\
&< x + 6k \\
&< n^2 - f(n) + 6n \\
&\leq n^2 - 4n \cdot g(3\sqrt{n}) \cdot \omega(n) \cdot \big(g(3\sqrt{n}) \cdot \omega(n) + 6\sqrt{n}\big) \\
&< n^2 - 4n \cdot g(3\sqrt{n}) \cdot \omega(n) \cdot \big(g(3\sqrt{n}) \cdot \omega(n) + 6\sqrt{n}\big) + \\
&\quad + \Big(2 \cdot g(3\sqrt{n}) \cdot \omega(n) \cdot \big(g(3\sqrt{n}) \cdot \omega(n) + 6\sqrt{n}\big)\Big)^2,
\end{aligned}$$

and so

$$k + 1 < n - 2 \cdot g(3\sqrt{n}) \cdot \omega(n) \cdot \big(g(3\sqrt{n}) \cdot \omega(n) + 6\sqrt{n}\big). \tag{2}$$

Then we have:

$$e_k\big(\mathrm{next}_S(s)\big) - b_k(s) = -\mathrm{next}_S(s)^2 + n - (k+1) + \mathrm{next}_S(s) + s^2 \\ > n - (k+1) + s^2 - \mathrm{next}_S(s)^2. \tag{3}$$

Using (1) we obtain:

$$e_k\big(\mathrm{next}_S(s)\big) - b_k(s) > n - (k+1) + s^2 - \Big(s + \big(\omega(k+1)+1\big)g(s_{max})\Big)^2 \\ = n - (k+1) - g(s_{max}) \cdot \big(\omega(k+1)+1\big)\Big(g(s_{max}) \cdot \big(\omega(k+1)+1\big) + 2s\Big)$$

using (2) we obtain:

$$\geq n - \Big(n - 2 \cdot g(3\sqrt{n}) \cdot \omega(n) \cdot \big(g(3\sqrt{n}) \cdot \omega(n) + 6\sqrt{n}\big)\Big) \\ - g(s_{max}) \cdot \big(\omega(k+1)+1\big) \cdot \Big(g(s_{max}) \cdot \big(\omega(k+1)+1\big) + 2s\Big) \\ \geq g(3\sqrt{n}) \cdot \sqrt{2} \cdot \omega(n) \cdot \big(g(3\sqrt{n}) \cdot \sqrt{2} \cdot \omega(n) + 6\sqrt{n}\big) \\ - g(s_{max}) \cdot \big(\omega(k+1)+1\big) \cdot \Big(g(s_{max}) \cdot \big(\omega(k+1)+1\big) + 2s\Big) \tag{4}$$

Observe that $3\sqrt{n} > (2+\varepsilon)\sqrt{k-1} \geq s_{max}$ from Lemma 11. Also, since $n > k+1$ (by (2)), for $n \geq 30$ we have $\omega(n) \geq 3$ and so $\sqrt{2} \cdot \omega(n) \geq \omega(k+1)+1$. Because functions g and ω are monotonic, we finally obtain that (4) ≥ 0.

Consider $e_k\big(\mathrm{next}_S(0)\big) - k^2$. Then it is equal to (3) with $s = 0$. All the above inequalities remains unchanged and then $k^2 \leq e_k(\mathrm{next}_S(0))$. ◀

Summarizing, for each $x \in [n^2/3, n^2 - f(n)]$ we find some k such that $x \in [(k-2)^2, k^2]$. From Lemma 11 we know that there exists $s_{max} \in S$ such that $b_k(s_{max}) \leq (k-2)^2$. From Lemma 12 we also get $k^2 \leq e_k\big(\mathrm{next}_S(0)\big)$, and the intervals for consecutive values from S between $\mathrm{next}_S(0)$ and s_{max} overlap. Hence, in particular, x belongs to some interval $I_k(s_x)$ where $s_x \in S$, $\mathrm{next}_S(0) \leq s_x \leq s_{max}$. Lemma 10 ensures that the pair $(k+1-s_x, k+1+s_x)$ belongs to $L(n)$ from the Theorem 6, so x is covered.

4 Automata with a sink state

It is known that if a synchronizing automaton $\mathscr{A}_n$ has a sink state, then $\mathrm{rt}(\mathscr{A}_n) \leq \frac{(n-1)n}{2}$ [26]. We show that reset thresholds of automata with a sink state cover all values in $[1, \dots, \frac{(n-1)n}{2}]$, at least if the size of the alphabet can be quadratic in n.

We construct a class of automata as follows. Let $T_n(V, E)$ be an undirected tree with $n = |V|$ vertices and the root r. We construct $\mathscr{A}(T_n) = (V, E, \delta)$, and the action of every letter $\{v_1, v_2\}$ is defined as follows:

$$\delta(v, \{v_1, v_2\}) = \begin{cases} v_1 & \text{if } v = v_2 \text{ and } r \notin \{v_1, v_2\}, \\ v_2 & \text{if } v = v_1 \text{ and } r \notin \{v_1, v_2\}, \\ r & \text{if } v, r \in \{v_1, v_2\}, \\ v & \text{otherwise.} \end{cases}$$

For $v \in V$ let $d(v)$ denote the distance in T_n from v to the root; hence $d(r) = 0$. For a subset $S \subseteq V$ we define:

$$U(T_n, S) = \sum_{q \in S} d(q).$$

▶ **Lemma 13.** *For every rooted tree $T_n(V,E)$ and every subset $S \subseteq V$ we have:* $\mathrm{rt}(\mathscr{A}(T_n), S) = U(T_n, S)$.

Proof. The proof follows easily by induction on $U(T_n,S)$. The case $U(T_n,S) = 0$ is trivial, since it must be that $S = \{r\}$. Let $U(T_n,S) \geq 1$ and assume that the claim holds for all S' such that $U(T_n,S') < U(T_n,S)$. Then observe that the action of any letter $\{v_1,v_2\} \in E$ increases the value of $U(T_n,S)$ by 1, decreases the $U(T_n,S)$ by 1, or does not change it. Moreover, we can always find such $\{v_1,v_2\}$ that decreases the value by 1, i.e. when $v_1 \in S$ and $v_2 \notin S$, or $v_1 = r$ and $v_2 \in S$. Thus the claim follows inductively and the automaton is synchronizing. ◀

▶ **Proposition 14.** *Given $n \geq 2$, for every $k \in [1,\ldots,\frac{(n-1)n}{2}]$ there exists an automaton $\mathscr{A}_n$ with a sink state and $\mathrm{rt}(\mathscr{A}) = k$.*

Proof. First, we consider the case of $k \geq n-1$. For each such k we construct a tree $T_n(V,E)$ with $U(T_n,V) = k$ and the claimed result follows by Lemma 13. We proceed by induction:

1. the *star* ($E = \{\{v,r\} \mid v \in V\}$) is suitable for the case of $n-1$;
2. let $T_n(V,E)$ be a tree such that $U(T_n,V) = k$. If the height of T_n is equal to $n-1$, then T_n is a *path* graph and $U(T_n,V) = \frac{(n-1)n}{2}$. Thus, we can assume that the height of T_n is smaller than $n-1$. We construct a new tree $T'_n(V',E)$ such that $U(T_n,V') = k+1$ in the following manner. Observe, that there exist two vertices v_1,v_2, both distinct from the root, that are leaves; without loss of generality let $d(v_1) \leq d(v_2)$. We will remove v_1, which does not change $d(v_2)$. Afterwards, we can find a vertex v_3 such that $d(v_3) = d(v_1)$, and attach v'_1 to v_3 as a leaf; thus $d(v'_1) = d(v_2)+1$ and $U(T_n,V') = k+1$.

Hence, we can construct the tree for any k starting from $n-1$.

Finally, for $k \in \{1,\ldots,n-2\}$ let $\mathscr{A}_n(V,E,\delta)$ be the automaton with $V = \{r,v_1,\ldots,v_{n-1}\}$, $E = \{e_1,\ldots,e_k\}$ and the transitions defined by

$$\delta(v_i,e_j) = \begin{cases} r & \text{if } i = j \text{ or } i > k, \\ v_i & \text{otherwise.} \end{cases}$$

Obviously, to synchronize $\mathscr{A}_n$ it is sufficient and necessary to use a word with every letter from E. ◀

5 General case

Let $\mathscr{C}_m(Q,\Sigma,\delta_{\mathscr{C}})$ be the Černý automaton with m states [10]: $Q = \{q_0,\ldots,q_{m-1}\}$, $\Sigma = \{a,b\}$, $\delta_{\mathscr{C}}(q_{m-1},a) = q_0$ and $\delta_{\mathscr{C}}(q_i,a) = q_i$ for $i \leq m-2$, and $\delta_{\mathscr{C}}(q_i,b) = (q_i+1) \bmod m$.

Let $T_{m'}(V,E)$ be a tree with m' vertices and let $\mathscr{A}(T_{m'}) = (V,E,\delta_{\mathscr{A}})$ be the automaton from Section 4.

Given $\mathscr{C}_m$ and $\mathscr{A}(T_{m'})$ we construct the joint automaton $\mathscr{B}_n$ with $n = m+m'-1$ states. This is done by union of both automata, while identifying the sink state r of $\mathscr{A}(T_{m'})$ with state q_0 of $\mathscr{C}_m$ and extending the actions of the letters for the states of the other automaton to identity.

Formally, assume that the alphabets Σ and E are disjoint, r is the sink state of $\mathscr{A}(T_{m'})$, and let $\mathscr{B}_n = (Q \cup V \setminus \{r\}, \Sigma \cup E, \delta)$, where δ is defined as follows:

$$\delta(q,a) = \begin{cases} \delta_{\mathscr{C}}(q,a) & \text{if } q \in Q \text{ and } a \in \Sigma, \\ \delta_{\mathscr{A}}(q,a) & \text{if } q \in V \setminus \{r\},\ a \in E, \text{ and } \delta_{\mathscr{A}}(q,a) \neq r, \\ q_0 & \text{if } q \in V \setminus \{r\},\ a \in E, \text{ and } \delta_{\mathscr{A}}(q,a) = r, \\ q & \text{if } q \in Q \text{ and } a \in E, \text{ or } q \in V \setminus \{r\} \text{ and } a \in \Sigma. \end{cases}$$

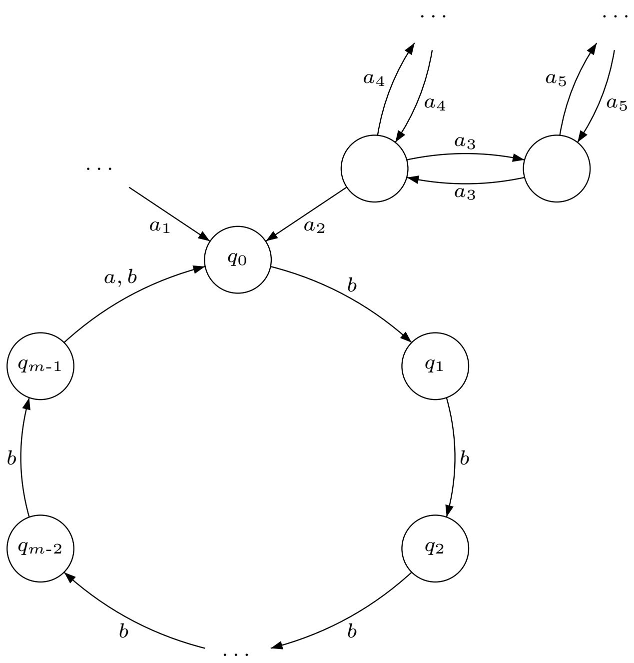

Figure 1 The automaton $\mathscr{B}_n$ from Section 5 formed from $\mathscr{C}_m$ and some $\mathscr{A}(T_{m'})$. The omitted transitions are self-loops.

The scheme of this construction is illustrated in Fig. 1.

▶ **Lemma 15.** *For all $m, m' \geq 1$ we have*

$$\operatorname{rt}(\mathscr{B}_n) = \operatorname{rt}(\mathscr{C}_m) + \operatorname{rt}(\mathscr{A}(T_{m'})).$$

Proof. Note that in the degenerated case $m' = 1$ this trivially holds, and suppose $m' \geq 2$.

To synchronize $\operatorname{rt}(\mathscr{B}_n)$ is it enough to apply the word $w'w$, where w' is a word synchronizing $\mathscr{A}(T_{m'})$ and w is a word synchronizing $\mathscr{C}_m$.

To show that this is a shortest possibility, let w be a synchronizing word for $\mathscr{B}_n$. Then w contain two disjoint subsequences of letters from Σ and from E, respectively. These subsequences synchronizes respectively $\mathscr{C}_m$ and $\mathscr{A}(T_{m'})$, which proves the lower bound. ◀

Arbitrary choice of m and m' with the constraint $n = m + m' - 1$ provides enough flexibility that finally leads to

▶ **Theorem 16.** *For every $k \leq n^2 - O(n^{3/2})$ there exists a synchronizing automaton with reset threshold k.*

Proof. Since $\operatorname{rt}(\mathscr{C}_m) = (m-1)^2$ and $\operatorname{rt}(\mathscr{A}(T_{m'})) \in \left[1, \ldots, \frac{(m'-1)m'}{2}\right]$ by Proposition 14, given m and m' we can construct an automaton with any value of reset threshold in $\left[(m-1)^2 + 1, (m-1)^2 + \frac{(m'-1)m'}{2}\right]$.

Since $n = m + m' - 1$ ($m' = n - m + 1$) we can cover all intervals

$$\left[(m-1)^2 + 1, (m-1)^2 + \frac{(n-m)(n-m+1)}{2}\right]$$

for all $1 \leq m \leq n$. Let g, $2 \leq g \leq (n-1)^2$, be the smallest number such that is not in the interval for some m. Consider the interval lying just before g ($g \geq 2$ so it exists), that is, let m

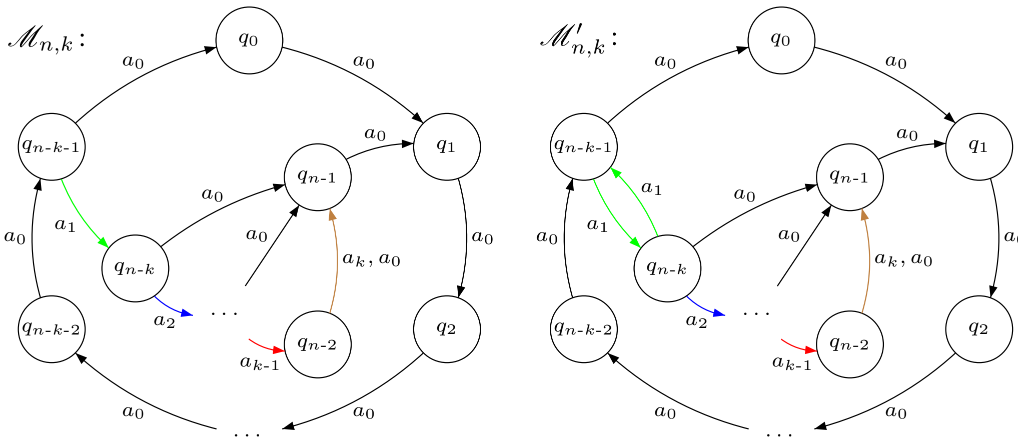

Figure 2 Automata $\mathscr{M}_{n,k}$ and $\mathscr{M}'_{n,k}$. The omitted transitions are self-loops.

be the largest number such that $g > (m-1)^2 + ((n-m)(n-m+1))/2$. Then the interval for $m+1$ must begin after g, so $g < m^2+1$. Hence $(m-1)^2 + ((n-m)(n-m+1))/2 + 1 \leq m^2$, which solved yields

$$m \geq \left(2n - \sqrt{16n+9} + 5\right)/2 = n - O(\sqrt{n}).$$

So $(m-1)^2 + (n-m)(n-m+1)/2 = n^2 - O(n^{3/2})$. ◀

6 Irreducibly synchronizing automata with large reset thresholds

Let $k \geq 1$ and $n \geq k+3$. Let $Q_n = q_0, \ldots, q_{n-1}$, and $\Sigma_k = a_0, \ldots, a_k$. We define the automaton $\mathscr{M}_{n,k} = (Q_n, \Sigma_k, \delta_{n,k})$, illustrated in Fig. 2 (left), with the transition function $\delta_{n,k}$ defined as follows:

$$\delta_{n,k}(q_i, a_0) = \begin{cases} q_{i+1} & \text{if } i \leq n-k-2 \\ q_0 & \text{if } i = n-k-1 \\ q_{n-1} & \text{if } n-k \leq i \leq n-2 \\ q_1 & \text{if } i = n-1, \end{cases}$$

and for $j \geq 1$

$$\delta_{n,k}(q_i, a_j) = \begin{cases} q_{i+1} & \text{if } i = n-k-2+j \\ q_i & \text{otherwise.} \end{cases}$$

We also define the variation $\mathscr{M}'_{n,k}(Q_n, \Sigma_k, \delta_{n,k})$, illustrated in Fig. 2 (right), of $\mathscr{M}_{n,k}$, where $\delta'_{n,k}$ is defined as follows:

$$\delta'_{n,k}(q_i, a_j) = \begin{cases} q_{n-k-1} & \text{if } i = n-k \text{ and } j = 1 \\ \delta_{n,k}(q_i, a_j) & \text{otherwise.} \end{cases}$$

▶ **Theorem 17.** *The automaton $\mathscr{M}_{n,k}$ is irreducibly synchronizing and has reset threshold*

$$n^2 - (k+3)n + 2k + 3.$$

Proof (sketch). The word $a_1 a_2 \cdots a_k (a_0^{n-k-1} a_1 \cdots a_k)^{n-k-2} a_0$ synchronizes $\mathscr{M}_{n,k}$ to the state q_1 and has length $n^2 - (k+3)n + 2k + 3$.

To show that this is the reset threshold, we use the backward tracing technique (cf. [19, Lemma 2], [21, 30]). The idea is to keep track, for $i = 1, 2, \ldots$, of families L_i of subsets of Q_n

that are preimages of a singleton under the action of a word of length i. Hence, L_i contains in particular all subsets that are compressible to a singleton by a word of length i. The smallest i such that $Q_n \in L_i$ is the length of the shortest reset words that synchronize to the singletons from L_0. Further, from each L_i we can exclude *visited* subsets, which are those being a proper subset of another subset from L_i or being a subset (non necessarily proper) of a subset from some $L_0, \ldots, L_{i-1}$ ([19, Lemma 2]). ◀

▶ **Theorem 18.** *The automaton $\mathscr{M}'_{n,k}$ is irreducibly synchronizing and has reset threshold*

$$n^2 - (k+3)n + 2k + 4.$$

References

1 D. Ananichev. A new lower bound for reset threshold of synchronizing automata with sink state. *ArXiv e-prints*, January 2017. `arXiv:1701.07954`.

2 D. S. Ananichev and V. V. Gusev. Approximation of reset thresholds with greedy algorithms. *Fundam. Inform.*, 145(3):221–227, 2016.

3 D. S. Ananichev, M. V. Volkov, and V. V. Gusev. Primitive digraphs with large exponents and slowly synchronizing automata. *Journal of Mathematical Sciences*, 192(3):263–278, 2013.

4 J. Araújo, P. J. Cameron, and B. Steinberg. Between primitive and 2-transitive: Synchronization and its friends. *ArXiv e-prints*, November 2015. `arXiv:1511.03184`.

5 R. C. Baker, G. Harman, and J. Pintz. The difference between consecutive primes, II. *Proceedings of the London Mathematical Society*, 83(3):532–562, 2001.

6 J. Berstel, D. Perrin, and C. Reutenauer. *Codes and Automata*. Encyclopedia of Mathematics an. Cambridge University Press, 2010.

7 V. D. Blondel, R. M. Jungers, and A. Olshevsky. On primitivity of sets of matrices. *Automatica*, 61(C):80–88, 2015.

8 R. A. Brualdi and Bolian Liu. Generalized exponents of primitive directed graphs. *Journal of Graph Theory*, 14(4):483–499, 1990.

9 R. A. Brualdi and H. Ryser. *Combinatorial Matrix Theory*. Cambridge University Press, 1991.

10 J. Černý. Poznámka k homogénnym experimentom s konečnými automatmi. *Matematicko-fyzikálny Časopis Slovenskej Akadémie Vied*, 14(3):208–216, 1964. In Slovak.

11 J. Černý, A. Pirická, and B. Rosenauerová. On directable automata. *Kybernetica*, 7:289–298, 1971.

12 H. Crámer. On the order of magnitude of the difference between consecutive prime numbers. *Acta Arithmetica*, 2:23–46, 1936.

13 P. Gawrychowski and D. Straszak. Strong inapproximability of the shortest reset word. In *Mathematical Foundations of Computer Science*, volume 9234 of *LNCS*, pages 243–255. Springer, 2015.

14 B. Gerencsér, V. V. Gusev, and R. M. Jungers. Primitive sets of nonnegative matrices and synchronizing automata. `https://arxiv.org/abs/1602.07556`, 2016.

15 V. Gusev. Lower Bounds for the Length of Reset Words in Eulerian Automata. *International Journal of Foundations of Computer Science*, 24(2), 2013.

16 V. V. Gusev and E. V. Pribavkina. Reset thresholds of automata with two cycle lengths. *International Journal of Foundations of Computer Science*, 26(07):953–966, 2015.

17 M. Ito and K. Shikishima-Tsuji. Some results on directable automata. In J. Karhumäki, H. Maurer, G. Păun, and G. Rozenberg, editors, *Theory Is Forever: Essays Dedicated to Arto Salomaa on the Occasion of His 70th Birthday*, pages 125–133. Springer, 2004.

18 J. Kari and M. V. Volkov. Černý's conjecture and the road coloring problem. In *Handbook of Automata*. European Science Foundation, 2013.

19 A. Kisielewicz, J. Kowalski, and M. Szykuła. Computing the shortest reset words of synchronizing automata. *Journal of Combinatorial Optimization*, 29(1):88–124, 2015.

20 A. Kisielewicz, J. Kowalski, and M. Szykuła. Experiments with Synchronizing Automata. In *Implementation and Application of Automata*, volume 9705 of *LNCS*, pages 176–188. Springer, 2016.

21 A. Kisielewicz and M. Szykuła. Synchronizing Automata with Extremal Properties. In *Mathematical Foundations of Computer Science*, volume 9234 of *LNCS*, pages 331–343. Springer, 2015.

22 Q. Li and J. Shao. The index set problem for boolean (or nonnegative) matrices. *Discrete Mathematics*, 123(1):75–92, 1993.

23 J. Olschewski and M. Ummels. The complexity of finding reset words in finite automata. In *Mathematical Foundations of Computer Science*, volume 6281 of *LNCS*, pages 568–579. Springer, 2010.

24 E. V. Pribavkina. Slowly synchronizing automata with zero and noncomplete sets. *Mathematical Notes*, 90(3):411, 2011.

25 G. Robin. Estimation de la fonction de Tchebychef ϕ sur le k-iéme nombre premier et grandes valeurs de la fonction $\omega(n)$ nombre de diviseurs premiers de n. *Acta Arithmetica*, 42(4):367–389, 1983.

26 I. K. Rystsov. Reset words for commutative and solvable automata. *Theoretical Computer Science*, 172(1-2):273–279, 1997.

27 J. Shao, J. Wang, and G. Li. Generalized primitive exponents with the complete characterizations of the extreme digraphs. *Chinese J. Contemp. Math.*, 15(4):317–324, 1994.

28 J. Shen and S. Neufeld. Local exponents of primitive digraphs. *Linear Algebra and its Applications*, 268:117–129, 1998.

29 M. Szykuła. Improving the upper bound on the length of the shortest reset words. *ArXiv e-prints*, February 2017. `arXiv:1702.05455`.

30 M. Szykuła and V. Vorel. An Extremal Series of Eulerian Synchronizing Automata. In *Developments in Language Theory*, LNCS, pages 380–392. Springer, 2016.

31 A. N. Trahtman. An efficient algorithm finds noticeable trends and examples concerning the Černý conjecture. In *Mathematical Foundations of Computer Science*, volume 4162 of *LNCS*, pages 789–800. Springer, 2006.

32 A. N. Trahtman. The Road Coloring Problem. *Israel Journal of Mathematics*, 172(1):51–60, 2009.

33 M. V. Volkov. Synchronizing automata and the Černý conjecture. In *Language and Automata Theory and Applications*, volume 5196 of *LNCS*, pages 11–27. Springer, 2008.