Algebraic power series and their automatic complexity

Eric Rowland
Hofstra University

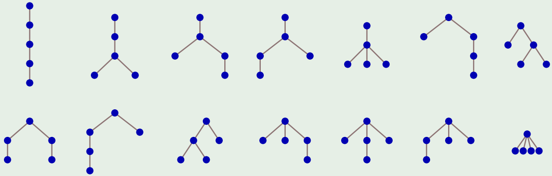
Joint work with Manon Stipulanti and Reem Yassawi

New York Combinatorics Seminar
CUNY Graduate Center, 2024–2–16

What do combinatorial sequences look like modulo p^{α}?

Example

Catalan numbers count plane trees with n edges:



$C(n)_{n\geq 0} = 1, 1, 2, 5, 14, 42, 132, 429, \ldots$

Modulo 2: $1, 1, 0, 1, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 1, \ldots$

$C(n)$ is odd if and only if $n+1$ is a power of 2.
(follows from Kummer 1852 since $C(n) = \frac{1}{n+1}\binom{2n}{n}$)

Modulo 4: $1, 1, 2, 1, 2, 2, 0, 1, 2, 2, 0, 2, 0, 0, 0, 1, \ldots$

Theorem (Eu–Liu–Yeh 2008)

For all $n \geq 0$,

$$C(n) \bmod 4 = \begin{cases} 1 & \text{if } n+1 = 2^a \text{ for some } a \geq 0 \\ 2 & \text{if } n+1 = 2^b + 2^a \text{ for some } b > a \geq 0 \\ 0 & \text{otherwise.} \end{cases}$$

In particular, $C(n) \not\equiv 3 \mod 4$.

Modulo 8: $1, 1, 2, 5, 6, 2, 4, 5, 6, 6, 4, 2, 4, 4, 0, 5, \ldots$

Theorem 4.2. *Let C_n be the nth Catalan number. First of all, $C_n \not\equiv_8 3$ and $C_n \not\equiv_8 7$ for any n. As for other congruences, we have*

$$C_n \equiv_8 \begin{cases} 1 & \text{if } n = 0 \text{ or } 1; \\ 2 & \text{if } n = 2^a + 2^{a+1} - 1 \text{ for some } a \geq 0; \\ 4 & \text{if } n = 2^a + 2^b + 2^c - 1 \text{ for some } c > b > a \geq 0; \\ 5 & \text{if } n = 2^a - 1 \text{ for some } a \geq 2; \\ 6 & \text{if } n = 2^a + 2^b - 1 \text{ for some } b - 2 \geq a \geq 0; \\ 0 & \text{otherwise.} \end{cases}$$

Liu and Yeh (2010) determined $C(n)$ mod 16:

Theorem 5.5. *Let c_n be the n-th Catalan number. First of all, $c_n \not\equiv_{16} 3, 7, 9, 11, 15$ for any n. As for the other congruences, we have*

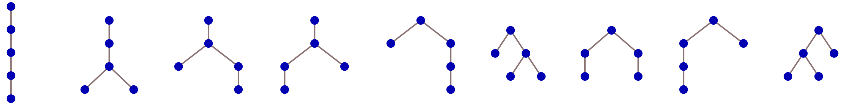
$$c_n \equiv_{16} \begin{cases} \left.\begin{matrix} 1 \\ 5 \\ 13 \end{matrix}\right\} & \text{if } \ d(\alpha) = 0 \ \text{ and } \begin{cases} \beta \le 1, \\ \beta = 2, \\ \beta \ge 3, \end{cases} \\ \left.\begin{matrix} 2 \\ 10 \end{matrix}\right\} & \text{if } \ d(\alpha) = 1,\ \alpha = 1 \ \text{ and } \begin{cases} \beta = 0 \text{ or } \beta \ge 2, \\ \beta = 1, \end{cases} \\ \left.\begin{matrix} 6 \\ 14 \end{matrix}\right\} & \text{if } \ d(\alpha) = 1,\ \alpha \ge 2 \ \text{ and } \begin{cases} (\alpha = 2, \beta \ge 2) \text{ or } (\alpha \ge 3, \beta \le 1), \\ (\alpha = 2, \beta \le 1) \text{ or } (\alpha \ge 3, \beta \ge 2), \end{cases} \\ \left.\begin{matrix} 4 \\ 12 \end{matrix}\right\} & \text{if } \ d(\alpha) = 2 \ \text{ and } \begin{cases} zr(\alpha) \equiv_2 0, \\ zr(\alpha) = 1, \end{cases} \\ 8 & \text{if } \ d(\alpha) = 3, \\ 0 & \text{if } \ d(\alpha) \ge 4. \end{cases}$$

where $\alpha = (CF_2(n+1) - 1)/2$ and $\beta = \omega_2(n+1)$ (or $\beta = \min\{i \mid n_i = 0\}$).

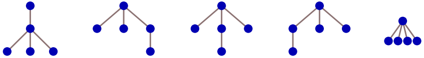
They also determined $C(n)$ mod 64.

What's the right framework?

Motzkin numbers count plane trees with n edges such that each vertex has at most 2 children:



Excluded:



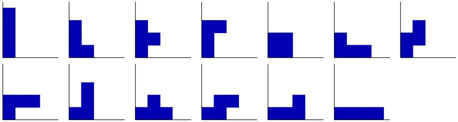
$M(n)_{n\geq 0} = 1, 1, 2, 4, 9, 21, 51, 127, \ldots$

Modulo 8: $1, 1, 2, 4, 1, 5, 3, 7, 3, 3, 4, 6, 7, 3, 2, 4, \ldots$

Theorem (Eu–Liu–Yeh; conj. by Deutsch–Sagan–Amdeberhan)

$M(n) \not\equiv 0 \mod 8$ *for all* $n \geq 0$.

Number of directed animals: $P(n)_{n \geq 0} = 1, 1, 2, 5, 13, 35, 96, 267, \ldots$



$P(4) = 13$

Number of restricted hexagonal polyominoes:
$H(n)_{n \geq 0} = 1, 1, 3, 10, 36, 137, 543, 2219, \ldots$

Riordan numbers: $R(n)_{n \geq 0} = 1, 0, 1, 1, 3, 6, 15, 36, \ldots$

Theorem (Deutsch–Sagan 2006)

There exists a set $C = \{1, 3, 4, 5, 7, \ldots\}$ with the property that

- *$P(n)$ is even if and only if $n \in 2C$,*
- *$H(n)$ is even if and only if $n \in 4C - 1$ or $n \in 4C$, and*
- *$R(n)$ is even if and only if $n \in 2C - 1$.*

Can we obtain and prove such results automatically?

Algebraic sequences

All these sequences $s(n)_{n\geq 0}$ are algebraic:

There is a nonzero polynomial $P(x, y)$ such that

$$P\left(x, \sum_{n\geq 0} s(n)x^n\right) = 0.$$

Example

For the Catalan numbers:

$\sum_{n\geq 0} C(n)x^n$ is a solution of $xy^2 - y + 1 = 0$ over $\mathbb{Q}$.

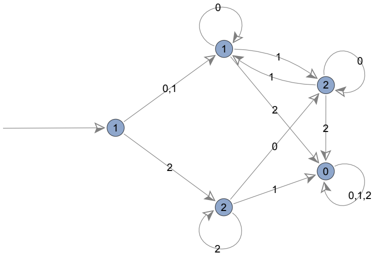
$\sum_{n\geq 0} (C(n) \bmod 3)x^n$ is a solution of $xy^2 + 2y + 1 = 0$ over $\mathbb{F}_3$.

$\mathbb{F}_p$ is the finite field with p elements.

Automatic sequences

$s(n)_{n\geq 0}$ is *p-automatic* if there is an automaton that outputs $s(n)$ when fed the base-p digits of n (least significant digit first).

$C(n) \bmod 3$:

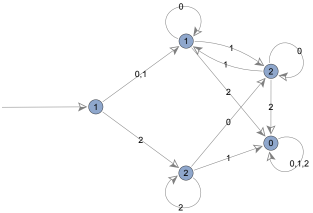


$C(9) \equiv ? \mod 3$. Since $9 = 100_3$, $C(9) \equiv \boxed{2} \mod 3$.

$(C(n) \bmod 3)_{n\geq 0} = 1, 1, 2, 2, 2, 0, 0, 0, 2, 2, \ldots$ is 3-automatic.

Two representations: polynomials and automata.

$$xy^2 + 2y + 1 = 0$$



Polynomial: easy to get from the polynomial over $\mathbb{Q}$.

Automaton: direct information about $s(n)$.

Theorem (Christol 1979/1980)

A sequence $s(n)_{n\geq 0}$ of elements in $\mathbb{F}_p$ is algebraic if and only if it is p-automatic.

How do we convert a polynomial into an automaton?

How does the automaton size depend on the polynomial degree?

How to tell whether a sequence is p-automatic?

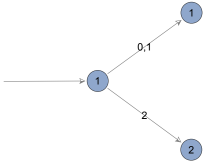
Let $r \in \{0, 1, \dots, p-1\}$.
The Cartier operator Λ_r picks out every pth term, starting with $s(r)$:

$$\Lambda_r(s(n)_{n\geq 0}) := s(pn+r)_{n\geq 0}$$

Iteratively apply $\Lambda_0, \Lambda_1, \dots, \Lambda_{p-1}$ to $s(n)_{n\geq 0}$.
Create one state in the automaton for each distinct sequence.

Let $s(n) = (C(n) \bmod 3)$. $\quad s(n)_{n\geq 0} = 1, 1, 2, 2, 2, 0, 0, 0, 2, \dots$.

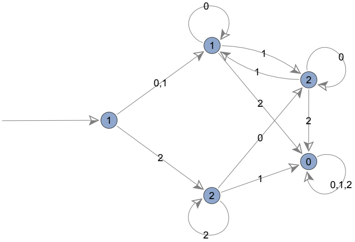
$\Lambda_0(s(n)_{n\geq 0}) = s(3n+0)_{n\geq 0} = 1, 2, 0, 2, 1, 0, 0, 0, 0, \dots$ new!
$\Lambda_1(s(n)_{n\geq 0}) = s(3n+1)_{n\geq 0} = 1, 2, 0, 2, 1, 0, 0, 0, 0, \dots = \Lambda_0(s(n)_{n\geq 0})$
$\Lambda_2(s(n)_{n\geq 0}) = s(3n+2)_{n\geq 0} = 2, 0, 2, 1, 0, 0, 0, 0, 2, \dots$ new!



Label each state with the initial term of the corresponding sequence.

$$\Lambda_0(\Lambda_0(s(n)_{n\geq 0})) = 1,2,0,2,1,0,0,0,0,2,\ldots = \Lambda_0(s(n)_{n\geq 0})$$
$$\Lambda_1(\Lambda_0(s(n)_{n\geq 0})) = 2,1,0,1,2,0,0,0,0,1,\ldots \qquad \text{new!}$$
$$\Lambda_2(\Lambda_0(s(n)_{n\geq 0})) = 0,0,0,0,0,0,0,0,0,0,\ldots \qquad \text{new!}$$

$\Lambda_r(\Lambda_2(s(n)_{n\geq 0}))$ ...



Eilenberg 1974:
A sequence is p-automatic if and only if this process terminates.

But we can't tell if sequences are equal from finitely many terms!

Use a different representation: diagonals of rational functions.

Polynomial for the Catalan numbers:

$$F = \sum_{n \geq 1} C(n) x^n \quad \text{satisfies} \quad x\,(F+1)^2 - F = 0.$$

Omit $C(0) = 1 \neq 0$.

Convert to the diagonal of a rational series (Furstenberg 1967):
$P = x(y+1)^2 - y$, so

$$F = \operatorname{diag}\left(\frac{y \frac{\partial P}{\partial y}(xy, y)}{P(xy, y)/y}\right) = \operatorname{diag}\left(\frac{y - 2xy^2 - 2xy^3}{1 - x - 2xy - xy^2}\right).$$

F mod 3 is the diagonal of

$$\begin{aligned}
\frac{y + xy^2 + xy^3}{1 + 2x + xy + 2xy^2} = \; & 0x^0y^0 + 1x^0y^1 + 0x^0y^2 + 0x^0y^3 + 0x^0y^4 + 0x^0y^5 + \cdots \\
& + 0x^1y^0 + 1x^1y^1 + 0x^1y^2 + 2x^1y^3 + 0x^1y^4 + 0x^1y^5 + \cdots \\
& + 0x^2y^0 + 1x^2y^1 + 2x^2y^2 + 0x^2y^3 + 1x^2y^4 + 2x^2y^5 + \cdots \\
& + 0x^3y^0 + 1x^3y^1 + 1x^3y^2 + 2x^3y^3 + 0x^3y^4 + 1x^3y^5 + \cdots \\
& + 0x^4y^0 + 1x^4y^1 + 0x^4y^2 + 2x^4y^3 + 2x^4y^4 + 0x^4y^5 + \cdots \\
& + 0x^5y^0 + 1x^5y^1 + 2x^5y^2 + 0x^5y^3 + 0x^5y^4 + 0x^5y^5 + \cdots \\
& + \cdots.
\end{aligned}$$

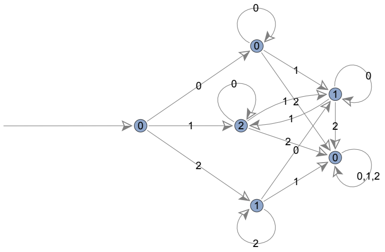
We have embedded $s(n)_{n\geq 1}$ into a series $\frac{S_0}{Q} := \frac{y+xy^2+xy^3}{1+2x+xy+2xy^2}$.

Construct an automaton by iterating $\lambda_{r,r}(S) := \Lambda_{r,r}\left(S \cdot Q^{p-1}\right)$.

$\lambda_{0,0}(S_0) = 2xy^2 + 2xy$ new!
$\lambda_{1,1}(S_0) = 1$ new!
$\lambda_{2,2}(S_0) = 2y + 2$ new!

$\lambda_{0,0}(2xy^2 + 2xy) = 2xy^2 + 2xy = \lambda_{0,0}(S_0)$ $\ldots$

Create one state in the automaton for each distinct polynomial.



The automaton may not be minimal.

Prime power moduli

This algorithm can be adapted to work modulo p^α.

Theorem (Denef–Lipshitz 1987)

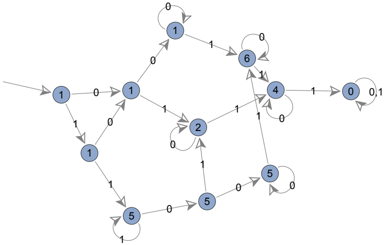
Let $\alpha \geq 1$. Let $R(\mathbf{x}), Q(\mathbf{x}) \in \mathbb{Z}_p[\mathbf{x}]$ such that $Q(0, \dots, 0) \not\equiv 0 \mod p$. Then the coefficient sequence of $\left(\operatorname{diag} \frac{R(\mathbf{x})}{Q(\mathbf{x})}\right) \mod p^\alpha$ is p-automatic.

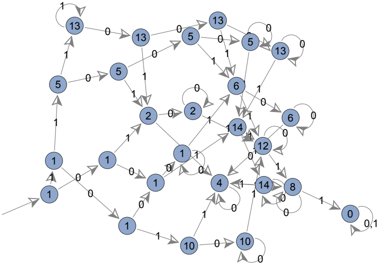
$\mathbb{Z}_p$ is the set of p-adic integers.

By computing an automaton for a sequence mod p^α, we can answer...

- Are there forbidden residues?
- What is the limiting distribution of residues (if it exists)?
- Is the sequence eventually periodic?
- Many other questions known to be decidable.

Catalan numbers modulo 8 and modulo 16:





Theorem (Liu–Yeh)

$C(n) \not\equiv 9 \mod 16$ *for all* $n \geq 0$.

Proof: Compute the automaton.

Catalan numbers modulo 2^α:

Theorem (Rowland–Yassawi 2015)

For all $n \geq 0$,

- $C(n) \not\equiv 17, 21, 26 \mod 32$,
- $C(n) \not\equiv 10, 13, 33, 37 \mod 64$,
- $C(n) \not\equiv 18, 54, 61, 65, 66, 69, 98, 106, 109 \mod 128$.

Only $\approx 35\%$ of the residues modulo 2^9 are attained by some $C(n)$.

Open question

Does the density of residues modulo 2^α attained by some Catalan number tend to 0 as α gets large?

Automaton size

How big is the (unminimized) automaton for $(C(n) \bmod 2^\alpha)_{n\geq 1}$?

α	1	2	3	4	5	6	7	8	9
size	4	6	15	37	83	194	445	1034	2403
$1.4 \times 2.3^\alpha$	3.2	7.4	17.0	39.2	90.1	207.3	476.7	1096.4	2521.6

height $h = \deg_x P$
degree $d = \deg_y P$

Upper bound from the construction: $p^{p^{2(\alpha-1)}\alpha h d}$

Example

$C(n) \bmod 2^9$: $P = x(y+1)^2 - y$ $h = 1$ $d = 2$
size $\leq 2^{18 \cdot 2^{16}} = 2^{1179648}$

Why is the bound so large?

Simpler setting: finite fields.

Theorem (Bridy 2017)

If the minimal polynomial P has height h and degree d, then the minimal automaton has size at most

$$(1 + o(1))\, p^{hd}$$

where $o(1)$ tends to 0 as any of p, h, d gets large.

Is the bound sharp? We suspect yes.

Polynomials in $\mathbb{F}_p[x, y]$ with maximum unminimized automaton size:

$p = 2$:

h	d	P	aut. size	p^{hd}	bound
1	2	$xy^2 + (x+1)y + x$	7	4	9
2	2	$x^2y^2 + (x^2+x+1)y + x^2$	14	16	25
3	2	$(x^3+x^2+1)y^2 + (x^3+1)y + x$	68	64	94
4	2	$(x^4+x+1)y^2 + (x^4+x^2+x+1)y + x$	252	256	311
5	2	$(x^5+x^3+1)y^2 + (x^5+x+1)y + x$	1052	1024	1192
6	2	$(x^6+x^5+1)y^2 + (x^6+x^2+x+1)y + x$	4062	4096	4424
7	2	$(x^7+x+1)y^2 + (x^7+x^4+x^3+x+1)y + x$	16424	16384	17288
1	3	$xy^3 + y^2 + (x+1)y + x$	11	8	18
2	3	$(x^2+x+1)y^3 + y^2 + (x^2+1)y + x^2 + x$	61	64	93
3	3	$(x^3+x+1)y^3 + y^2 + (x^3+x^2+x+1)y + x^3 + x^2$	533	512	614
4	3	$(x^4+x+1)y^3 + y^2 + (x^4+1)y + x^4 + x^3 + x$	4213	4096	4871
1	4	$(x+1)y^4 + y^2 + (x+1)y + x$	20	16	33
2	4	$(x^2+x+1)y^4 + y^3 + (x^2+x+1)y + x^2 + x$	216	256	358
3	4	$(x^3+x+1)y^4 + y^3 + (x^3+1)y + x^2 + x$	3956	4096	4870
1	5	$(x+1)y^5 + (x+1)y^2 + y + x$	37	32	67
2	5	$(x^2+x+1)y^5 + y^4 + y^3 + x^2y^2 + y + x^2 + x$	889	1024	1510
3	5	$(x^3+x^2+1)y^5 + y^4 + x^3y^2 + (x+1)y + x^3 + x^2 + x$	43913	32768	48134

$p = 3$:

h	d	P	aut. size	p^{hd}	bound
1	2	$(x+1)y^2 + y + x$	9	9	14
2	2	$(x^2+x+2)y^2 + y + x^2$	79	81	91
3	2	$(x^3+x^2+2x+1)y^2 + y + x^3 + x$	727	729	788
4	2	$(x^4+x^3+2)y^2 + y + x^4 + x$	6533	6561	6729

Can we get Bridy's bound without algebraic geometry? Yes.

Theorem (Rowland–Stipulanti–Yassawi 2023)

The minimal automaton has size at most

$$p^{hd} + p^{(h-1)(d-1)} L(h) L(d)^2 + \lfloor \log_p h \rfloor + \lceil \log_p \max(h, d-1) \rceil + 3.$$

$P \in \mathbb{F}_p[x, y], \quad h = \deg_x P, \quad d = \deg_y P$

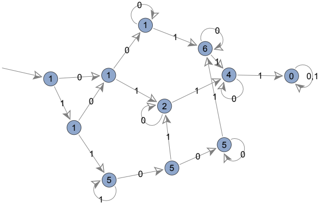
$L(n)$ is the Landau function:

$$L(5) = \max(\operatorname{lcm}(5), \operatorname{lcm}(4,1), \operatorname{lcm}(3,2), \operatorname{lcm}(3,1,1), \operatorname{lcm}(2,2,1), \operatorname{lcm}(2,1,1,1), \operatorname{lcm}(1,1,1,1,1)) = 6$$

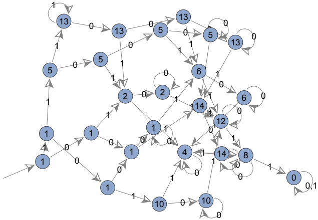
3 univariate polynomials R arise, with degrees $\leq h, d, d$.

Factor each $R = R_1^{e_1} \cdots R_k^{e_k}$. $\longrightarrow$ period length $\operatorname{lcm}(\deg R_1, \ldots, \deg R_k)$,
transient length $\log_p \max(e_1, \ldots, e_k)$

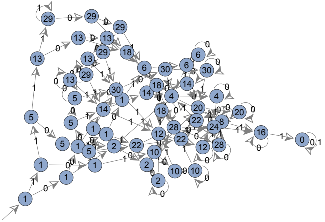
mod 8:



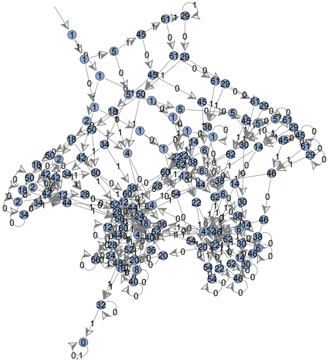
mod 16:



mod 32:



mod 64:



These automata project to each other.

So there is an inverse limit profinite automaton. Can we describe it?

Let $D = \{0, 1, \ldots, p-1\}$.

Theorem

Every state in the automaton for $(s(n) \bmod p^\alpha)_{n \geq 0}$ *is of the form*

$$T_0 Q^{p^{\alpha-1}-1} + p T_1 Q^{p^{\alpha-1}-2} + p^2 T_2 Q^{p^{\alpha-1}-3} + \cdots + p^{\alpha-1} T_{\alpha-1} Q^{p^{\alpha-1}-\alpha}$$

where $T_i \in D[x, y]$ *for each* $i \in \{0, 1, \ldots, \alpha - 1\}$.

Equivalently: $\left(T_0 + T_1 \frac{p}{Q} + T_2 \left(\frac{p}{Q}\right)^2 + \cdots + T_{\alpha-1} \left(\frac{p}{Q}\right)^{\alpha-1}\right) Q^{p^{\alpha-1}-1}$.

There are bounds on $\deg_x T_i$ and $\deg_y T_i$.

Singly exponential upper bound:

$$(1 + o(1))\, p^{\frac{1}{6}\alpha(\alpha+1)((2hd-1)\alpha+hd+1)}$$

When $\alpha = 1$, we recover Bridy's $(1 + o(1))\, p^{hd}$ for $\mathbb{F}_p$.

References 1

- Andrew Bridy, Automatic sequences and curves over finite fields, *Algebra & Number Theory* **11** (2017) 685–712.
- Gilles Christol, Teturo Kamae, Michel Mendès France, and Gérard Rauzy, Suites algébriques, automates et substitutions, *Bulletin de la Société Mathématique de France* **108** (1980) 401–419.
- Jan Denef and Leonard Lipshitz, Algebraic power series and diagonals, *Journal of Number Theory* **26** (1987) 46–67.
- Emeric Deutsch and Bruce E. Sagan, Congruences for Catalan and Motzkin numbers and related sequences, *Journal of Number Theory* **117** (2006) 191–215.
- Sen-Peng Eu, Shu-Chung Liu, and Yeong-Nan Yeh, Catalan and Motzkin numbers modulo 4 and 8, *European Journal of Combinatorics* **29** (2008) 1449–1466.

References 2

- Harry Furstenberg, Algebraic functions over finite fields, *Journal of Algebra* **7** (1967) 271–277.
- Shu-Chung Liu and Jean C.-C. Yeh, Catalan numbers modulo 2^k, *Journal of Integer Sequences* **13** (2010) Article 10.5.4 (26 pages).
- Eric Rowland and Reem Yassawi, Automatic congruences for diagonals of rational functions, *Journal de Théorie des Nombres de Bordeaux* **27** (2015) 245–288.
- Eric Rowland, Manon Stipulanti, and Reem Yassawi, Algebraic power series and their automatic complexity I: finite fields, `https://arxiv.org/abs/2308.10977`.