

WHERE THE SUBSTITUTION METHOD STOPS ON THE SHORT PRODUCT OVER $\mathbb{F}_2$

KOREA SUPERINTELLIGENCE LABS

ABSTRACT. Let T_N be the structure tensor of the short (truncated) product $\mathbb{F}_2[x]/x^N$, whose rank $R(T_N)$ is the number of coefficient multiplications a bilinear algorithm for that product needs. The value $R(T_5)$ sits at the boundary of what exhaustive search reaches: the bound $R(T_5) \geq 10$ is classical, the matching $R(T_5) \geq 11$ rests on a single covering-sets computation reported in a summary table, and no independent verification of it exists.

We show that the substitution method, in the form of the explicit one-level ladder it yields for T_{N+1} , provably cannot supply the missing unit. Restated on subspaces of matrices, the ladder reduces $R(T_5) \geq 3+m$ to the minimum m of a rank functional over a family of 4-dimensional subspaces of $\mathbb{F}_2^{4 \times 4}$ indexed by 4096 parameter triples. We prove that this minimum is *exactly* 7: one member is contained in the span of seven rank-one matrices – and exactly seven rank-one matrices lie in that span, so the witness is complete – while no member at all is contained in the span of six, a statement certified by exhausting 103,643,136 candidate subspaces. Since the slice space of T_4 needs eight, the comparison is exact and the ladder yields $R(T_5) \geq 10$, never 11.

The same machinery produces certificates, checked by the Lean 4 kernel, for the exact values $R(T_3) = 5$ and $R(T_4) = 8$ in both directions, and for the exact symmetric ranks of $\mathbb{F}_2[x]/x^N$ for $N \leq 5$ and of $\mathbb{F}_2[x]/(x^N - 1)$ for $N \in \{3, 5, 6\}$, together with $R^{\text{sym}}(T_6) \in \{13, 14\}$. Those values are published; the certificates are new.

1. INTRODUCTION

1.1. The object. Fix a commutative ring ρ and $N \geq 1$. The *short product* (or *truncated product*) of length N multiplies two polynomials of degree $< N$ over ρ and keeps the coefficients of $1, x, \dots, x^{N-1}$; it is the multiplication of the local algebra $\rho[x]/x^N$. Its structure tensor is

$$T_N[i][j][k] = [i + j = k], \quad 0 \leq i, j, k < N,$$

and over a field $\mathbb{F}$ the tensor rank $R_{\mathbb{F}}(T_N)$ is the *bilinear complexity* of $\mathbb{F}[x]/x^N$: the least number of coefficient multiplications a bilinear algorithm for the short product performs. Throughout, $\mathbb{F} = \mathbb{F}_2$ and we write R for $R_{\mathbb{F}_2}$. We also consider the cyclic product C_N , the structure tensor of $\mathbb{F}_2[x]/(x^N - 1)$, obtained by reading the index condition as $i + j \equiv k \pmod{N}$.

The short product is the recurring base case of Karatsuba- and Toom-style multiplication schemes, and the exact values of $R(T_N)$ have been pursued for decades. Over $\mathbb{F}_2$, $R(T_2) = 3$ is Karatsuba's, and $R(T_3) = 5$ and $R(T_4) = 8$ are classical: Table 6 of [13] credits the matching lower bounds to Winograd [14] with Alder–Strassen [1] at $N = 3$, and to Averbuch–Galil–Winograd [2] with Bläser [5] at

Date: August 30, 2026.

$N = 4$. Their optimality over $\mathbb{F}_2$ was re-proved by exhaustive search by Barbulescu–Detrey–Estibals–Zimmermann [4], whose §5 states that their computation “*proves the tensor rank from [18] is the optimal one for $n = 2, 3, 4$ over $\mathbb{F}_2$ modulo both X^n and $X^n - 1$* ” — their reference [18] being Oseledets [10].

1.2. The cell this paper is about. At $N = 5$ the record is thinner than it looks. Row $N = 5$ of the $\mathbb{F}_2$ columns of Table 6 of the recent automated-lower-bound work of Wang [13] reads: previous lower bound 10, credited to Averbuch–Galil–Winograd [2] and Bläser [5]; that paper’s own lower bound 10; previous upper bound 11, credited to Cenk–Özbudak [7]. The exhaustive search of [4] closes neither end: its Table 4 runs the general case at $n = 5$ only up to $k = 9$ (empty, $2.66 \cdot 10^{10}$ tests) and never runs $k = 10$, and its optimality claim is made only for $n = 2, 3, 4$. That is where the bracket $R(T_5) \in [10, 11]$ comes from, and it is the bracket still displayed — as a one-wide gap — in that table, whose own new lower bounds in the $\mathbb{F}_2$ columns begin at $N = 6$.

We did not consult [2, 5, 7] directly; the attributions just quoted are Wang’s, read from the v10 e-print source of [13] and from its own compiled table numbering, and the bibliographic data of those three items were verified against Crossref.

The gap is closed in the literature, by a source that Wang’s reference list does not contain: Covanov [8] reports $\text{rank}(\text{ShortProduct}_5) = 11$ over $\mathbb{F}_2$ and enumerates all 146,944 optimal decompositions in 11,424 equivalence classes, obtained by the covering-sets method of that paper. So $R(T_5) = 11$; but the lower half of that equality rests on a single unverified exhaustion, reported in a summary table rather than as a numbered theorem, and no independent verification of $R(T_5) \geq 11$ has appeared.

Two natural questions follow, and this paper answers the first.

- *Can the substitution method — the cheap, general lower-bound technique, and the one Wang automates — reach 11?* It reaches 10. Whether that is a limitation of the implementations or of the method has not been addressed; [13] reports the bound 10 at T_5 over $\mathbb{F}_2$, and the only cells it discusses are the five where its own bound falls one below the previously known one, which this cell is not.
- *Can the exhaustion behind $R(T_5) \geq 11$ be independently re-run?* Not by a naive complete search: we measure the cost below at $6.8 \cdot 10^{12}$ candidate subspaces, about $2 \cdot 10^4$ core-years.

1.3. Results. Our main result is that the substitution method has a ceiling here, and that the ceiling is exactly the classical bound. We work with the standard reformulation of tensor rank as a minimum over subspaces of matrices (§2): for a subspace $V \subseteq \mathbb{F}_2^{N \times N}$ and a set G of *atoms*,

$$R_G(V) = \min \{ \dim W : V \subseteq W \subseteq \mathbb{F}_2^{N \times N}, W \text{ spanned by the atoms in it } \},$$

with G the rank-one matrices for ordinary rank and the symmetric rank-one matrices uu^T for the symmetric restriction. The substitution method, applied once to T_{N+1} , yields the ladder inequality

$$(\star) \quad R(T_{N+1}) \geq 3 + \min_{\lambda, \mu, \nu \in \mathbb{F}_2^N} R(V(\lambda, \mu, \nu))$$

for an explicit family $V(\lambda, \mu, \nu)$ of N -dimensional subspaces of $\mathbb{F}_2^{N \times N}$ (Proposition 6).

- **Theorem 8:** at $N = 4$ the member of the family at $(\lambda, \mu, \nu) = (e_2, e_2, e_2 + e_3)$ is contained in the span of seven rank-one matrices, and *exactly* seven rank-one matrices lie in that span — so no larger admissible space hides inside it. Against it, the slice space of T_4 is contained in the span of no seven rank-one matrices.
- **Theorem 9:** no member of the family at $N = 4$, over all 4096 parameter triples, is contained in the span of six rank-one matrices. Hence the minimum in $(\star)$ is exactly 7, and one level of the ladder yields exactly $R(T_5) \geq 10$ for every parameter choice: the obstruction is not one unlucky triple.

The comparison in Theorem 8 is exact rather than analogical: both halves are instances of one predicate, “contained in the span of at most r rank-one matrices”, and both are decided by the same program. The conclusion is that $R(T_5) \geq 11$ needs covering sets or an equivalent global technique — the missing ingredient is the algorithm, not the compute.

Around that result we place certificates for the values the argument uses and for their symmetric analogues. All of these values are published; what is new is that each is now carried by a machine-checkable artifact, which none of the sources we consulted ships — §6.1 states the scope of that claim, which rests on our own literature search.

- $R(T_2) = 3$, $R(T_3) = 5$, $R(T_4) = 8$, both directions certified (Theorems 14 and 15); the values are classical and their optimality over $\mathbb{F}_2$ is [4].
- The exact symmetric ranks $R^{\text{sym}}(T_2) = 3$, $R^{\text{sym}}(T_3) = 5$, $R^{\text{sym}}(T_4) = 8$, $R^{\text{sym}}(T_5) = 11$, $R^{\text{sym}}(C_3) = 4$, $R^{\text{sym}}(C_5) = 10$, $R^{\text{sym}}(C_6) = 12$, and the bracket $R^{\text{sym}}(T_6) \in \{13, 14\}$ (§6). The values are the (Sym.) rows of Table 4 of [4], four of them also printed as exact values in Rambaud’s thesis [12].
- Explicit checked bilinear algorithms realising the upper ends: $R(T_5) \leq 11$, $R(T_6) \leq 14$, $R(C_3) \leq 4$, $R(C_5) \leq 10$, $R(C_6) \leq 12$ (§5.1).

Two of these speak to cells that are still open. Wang [13] brackets $R(T_6) \in [13, 14]$ and $R(C_6) \in [11, 12]$; our $R^{\text{sym}}(C_6) = 12$ says that a rank-11 scheme for C_6 , if one exists, is necessarily non-symmetric, and $R^{\text{sym}}(T_6) \geq 13$ is the corresponding half-statement for T_6 .

Every theorem, proposition and lemma below has been verified in Lean 4 (§8); the single exception is Proposition 6, which is a hand proof and is flagged as such. The remarks are prose, and the one that reports measurements rather than mathematics says so in its heading.

2. PRELIMINARIES

2.1. Schemes and identities. The definitions are the standard ones for bilinear complexity; see [6, Ch. 14]. Let ρ be a commutative semiring.

Definition 1. A *rank- R scheme* for the length- N short product over ρ is a triple of coefficient tables $a, b, c: \{0, \dots, R-1\} \times \{0, \dots, N-1\} \rightarrow \rho$. It is *correct* when it satisfies the N^3 *defining identities*

$$\sum_{r < R} a_{ri} b_{rj} c_{rk} = [i + j = k], \quad 0 \leq i, j, k < N.$$

We write $R_\rho(T_N) \leq R$ when a correct rank- R scheme exists.

The scheme describes an algorithm: form the R products $P_r = L_r(A)L'_r(B)$, where $L_r(A) = \sum_i a_{ri}A_i$ and $L'_r(B) = \sum_j b_{rj}B_j$, then read off the k -th output coefficient as $\sum_r P_r c_{rk}$. That this algorithm is correct exactly when the identities hold is the standard trilinear computation, and it is worth isolating because it mentions no particular scheme: it is proved once, and every finite per-scheme check inherits it.

Proposition 2 (Soundness). *Let ρ be a commutative semiring and (a, b, c) a rank- R scheme satisfying the N^3 defining identities. Then for all $A, B \in \rho^N$ and all $k < N$,*

$$\sum_{r < R} \left(\sum_{i < N} a_{ri} A_i \right) \left(\sum_{j < N} b_{rj} B_j \right) c_{rk} = \sum_{i+j=k} A_i B_j,$$

i.e. the scheme computes the short product coefficientwise.

Proposition 3 (Transfer and padding). *Let $f: \rho \rightarrow \sigma$ be a homomorphism of commutative semirings. If (a, b, c) satisfies the identities over ρ , then (fa, fb, fc) satisfies them over σ ; hence $R_\rho(T_N) \leq R$ implies $R_\sigma(T_N) \leq R$. In particular a scheme over $\mathbb{F}_2$ is a scheme over every commutative $\mathbb{F}_2$ -algebra, and a scheme over $\mathbb{Z}$ is a scheme over every commutative ring. Moreover $R_\rho(T_N) \leq R$ implies $R_\rho(T_N) \leq R'$ for every $R' \geq R$, by padding with zero products.*

The identities are polynomial in the table entries with 0/1 right-hand sides, which is the whole content of the transfer statement — and what makes a single finite check over $\mathbb{F}_2$ a statement about every commutative $\mathbb{F}_2$ -algebra.

Lemma 4 (The decidable layer). *For tables over a commutative semiring with decidable equality, the Boolean check that evaluates all N^3 identities is equivalent to the identities themselves. Table lookup is by index with default 0, so an out-of-range entry reads as 0 and the check and the statement see the same scheme.*

Lemma 4 is what lets a *failing* check refute a claim, and it is used for the negative controls of §5.3.

2.2. The slice space and the atom reformulation. Let $M_k \in \mathbb{F}_2^{N \times N}$ be the k -th mode-3 slice of T_N , that is $M_k[i][j] = [i + j = k]$, and let $V_N = \langle M_0, \dots, M_{N-1} \rangle$ be the *slice space*, of dimension N . For a set G of matrices — the *atoms* — call a subspace W *G-admissible* if W is spanned by the atoms it contains, and put

$$R_G(V) = \min\{\dim W : V \subseteq W, W \text{ G-admissible}\}.$$

Taking $G = \{uv^\top : u, v \neq 0\}$, the rank-one matrices, gives $R_G(V_N) = R(T_N)$; taking $G = \{uu^\top : u \neq 0\}$ gives the least rank of a *symmetric* scheme. We abbreviate $R(V) = R_G(V)$ for G the rank-one matrices. The reformulation is classical — it is the search space of Algorithm 1 of [4] — and both directions are one line: a rank-one basis of W expands each slice into a decomposition of that length, and conversely the r rank-one summands of a decomposition span a G -admissible $W \supseteq V$ of dimension at most r . Only the second direction is used below, and only through Theorem 5.

2.3. Symmetric schemes. A scheme is *symmetric* when $b = a$, i.e. when the decomposition has the shape $T = \sum_t a_t \otimes a_t \otimes c_t$. We write R^{sym} for the least rank of a symmetric scheme; this is the (Sym.) restriction of [4, §4.6], and the quantity denoted μ^{sym} in [12]. Dumas, Lia and Sheekey [9], adapting the same search to

semifields, say of decompositions with $L = R$: “We refer to this as the *partially symmetric tensor rank*”; R^{sym} is that quantity. Trivially $R \leq R^{\text{sym}}$.

One caveat travels with the restriction. As [4, §4.6] points out, symmetric rank-one atoms may fail to produce optimal formulae for a symmetric map — their example is $a_0b_1 + a_1b_0$, which cannot be computed with two symmetric products over $\mathbb{F}_2$ — so an exhaustion over symmetric atoms alone bounds R^{sym} from below and says nothing about R . Every symmetric statement in this paper is accordingly a statement about R^{sym} .

3. THE SPAN ENGINE

All exhaustive statements in this paper are answers of one program, and the program is certified by one theorem. We describe both, because the credibility of the negative results rests on them.

Over $\mathbb{F}_2$ a matrix is a bit vector and addition is exclusive-or, so a subspace is presented as an *echelon list* of (pivot, vector) pairs, maintained so that each vector carries its own pivot and no later vector has a 1 there. Reduction against such a list is the membership test. The search walks the atom list, adjoining an atom only when it is independent of the current span, at most b times, and answers **true** as soon as the atoms lying inside the current W already span W . Started at an echelon basis of V with budget $b = r - \dim V$, an answer of **false** certifies $R_G(V) > r$.

Theorem 5 (Correctness of the engine). *Let atoms be a list of matrices, V a list of matrices, and $r \in \mathbb{N}$. If some list $g_1, \dots, g_s$ with $s \leq r$, each g_t either zero or a member of atoms, has every member of V in its $\mathbb{F}_2$ -span, then the search started at an echelon basis of V with budget $r - \dim V$ answers **true**.*

Contrapositively: an answer of false rules out every such list, hence every rank- r decomposition of the tensor whose slice space is spanned by V into products drawn from atoms.

Two facts carry the proof, and both are stated for echelon lists over $\mathbb{F}_2$. First, reduction against an echelon list is $\mathbb{F}_2$ -linear with kernel exactly the span, so the *computed* membership test is the *mathematical* one. Second, an echelon list of length d has exactly 2^d elements in its span — the selection map from $\{0, \dots, 2^d - 1\}$ is always surjective onto the span and is injective exactly under the pivot invariant — so two echelon lists with the same span have the same length. The second fact replaces the Steinitz exchange lemma, which is the step that completeness of the search would otherwise need, by a cardinality count; over $\mathbb{F}_2$ this is much the shorter route. Completeness itself is then an induction over sublists of the atom list whose only idea is that skipping an atom already inside the current span loses nothing.

The bridge from schemes to the engine is the observation that a rank- R scheme over $\mathbb{F}_2$ presents slice k of the tensor as the exclusive-or of the rank-one matrices $a_r b_r^\top$ over those r with $c_{rk} = 1$ — which is the system of N^3 defining identities, read one matrix entry at a time. Combining this with Theorem 5 gives the two statements every lower bound below instantiates: if the general (resp. symmetric) search returns **false** at rank r , then no scheme (resp. no symmetric scheme) of rank r exists.

4. THE LADDER AND ITS CEILING

4.1. The substitution ladder. The substitution method, restated on subspaces of matrices, is three moves, each costing one dimension. Let $V \subseteq \mathbb{F}^{p \times q}$ be a subspace over an arbitrary field $\mathbb{F}$.

- *Row fold.* Let $r < p$ and suppose some $M \in V$ has nonzero row r . For $\lambda \in \mathbb{F}^{p-1}$ let $\Phi_{r,\lambda}$ delete row r after adding $\lambda_i \cdot (\text{row } r)$ to each surviving row i . Then $R(V) \geq 1 + \min_{\lambda} R(\Phi_{r,\lambda} V)$.
- *Column fold.* The transpose statement.
- *Slice kill.* Let $0 \neq M \in V$ and $V = V_0 \oplus \langle M \rangle$. Then $R(V) \geq 1 + \min_{\psi \in V_0^*} R(\{x + \psi(x)M : x \in V_0\})$.

Each proof is two lines from the reformulation: take a rank-one basis $w_1, \dots, w_d$ of an optimal admissible $W \supseteq V$; some $w_t = uv^\top$ has $u_r \neq 0$, and normalising $u_r = 1$ and setting $\lambda_i = -u_i$ sends w_t to zero while keeping every other image of rank at most one, so the image space is spanned by at most $d - 1$ rank-one matrices. The slice kill is the same argument with the roles of the basis played by the coefficients of M . The prover chooses the move; the adversary chooses the parameter.

Applying the row fold to row N of T_{N+1} , then the column fold to column N , then the slice kill to the slice they create, and computing the three images, gives a closed form.

Proposition 6. *For every field $\mathbb{F}$ and every $N \geq 2$,*

$$(\star) \quad R(T_{N+1}) \geq 3 + \min_{\lambda, \mu, \nu \in \mathbb{F}^N} R(\langle M_k + \nu_k C(\lambda, \mu) : 0 \leq k < N \rangle),$$

where $M_k[i][j] = [i + j = k]$ and $C(\lambda, \mu)[i][j] = [i + j = N] + \lambda_i[j = 0] + \mu_j[i = 0]$ for $i, j < N$. The three side conditions hold for $N \geq 2$: row N of the last slice of T_{N+1} is e_0 ; after the row fold, column N is e_0 ; and $C(\lambda, \mu)$ is supported partly on $i + j = N$, outside the support of the M_k , so it is not in their span.

Remark 7. Proposition 6 is the only statement in this paper proved by hand: it is not part of the machine-checked development. Everything the ceiling argument adds to it is machine-checked.

On attribution: the substitution method is textbook material — [6, Ch. 6] is titled *The Substitution Method* — and [13] attributes it to Pan [11]; the version automated there is that paper’s Lemma 3 (v10), stated on linear forms rather than on subspaces of matrices. We did not find the three-move matrix-subspace formulation above stated in that shape in any of them, so we claim no more for it than that it is a routine restatement, and the two-line proofs above derive each move from the reformulation of §2 rather than citing one.

Write $V(\lambda, \mu, \nu) = \langle M_k + \nu_k C(\lambda, \mu) : k < N \rangle$ for the family in $(\star)$. Over $\mathbb{F}_2$ at $N = 4$ it has 4096 members, one per parameter triple in $(\mathbb{F}_2^4)^3$.

4.2. The obstruction. We pack a matrix in $\mathbb{F}_2^{4 \times 4}$ row-major, $E_{ij} \mapsto \text{bit } 4i + j$, and write vectors of $\mathbb{F}_2^4$ in the basis $e_0, \dots, e_3$.

Theorem 8. *Let $(\lambda, \mu, \nu) = (e_2, e_2, e_2 + e_3)$. Then $C(\lambda, \mu) = E_{02} + E_{13} + E_{20} + E_{22} + E_{31}$ and*

$$V(\lambda, \mu, \nu) = \left\langle \begin{matrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{matrix}, \begin{matrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{matrix}, \begin{matrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{matrix}, \begin{matrix} 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 1 & 1 & 1 & 0 \\ 1 & 1 & 0 & 0 \end{matrix} \right\rangle,$$

a 4-dimensional space. Let

$$G_0 = \{uu^\top : u \in \{e_0, e_2, e_0+e_1+e_2, e_3, e_0+e_3, e_1+e_3, e_0+e_1+e_3\}\},$$

a set of seven rank-one matrices. Then:

- (1) $V(\lambda, \mu, \nu) \subseteq \langle G_0 \rangle$, and $\dim \langle G_0 \rangle = 7$;
- (2) the rank-one matrices lying in $\langle G_0 \rangle$ are exactly the seven members of G_0 , so no eighth rank-one matrix and hence no larger admissible subspace hides inside $\langle G_0 \rangle$;
- (3) the slice space V_4 of T_4 is contained in the span of no seven rank-one matrices.

Consequently $R(V(\lambda, \mu, \nu)) \leq 7 < 8 = R(T_4)$, the last equality being Theorem 15.

Proof sketch. Parts (1) and (2) are finite checks over a 16-bit space: (1) reduces the four generators against an echelon basis of $\langle G_0 \rangle$ and observes that $\dim \langle G_0 \rangle = 7$; (2) filters the 225 rank-one matrices $uv^\top$, $u, v \neq 0$, in $\mathbb{F}_2^{4 \times 4}$ by membership in $\langle G_0 \rangle$ and finds the listed seven. Both are decided by evaluation inside the proof kernel. The displayed generators of $V(\lambda, \mu, \nu)$ are likewise obtained by evaluating the definition of the family at the stated parameters, not by transcription.

Part (3) is the contrapositive of Theorem 5 applied to the answer **false** of the general search at $N = 4$, $r = 7$. That search is exhaustive: it walks every subspace $W \supseteq V_4$ obtained by adjoining at most three of the 225 rank-one matrices, testing each for admissibility, and rejects all of them — 1,867,373 candidate subspaces. The count was reproduced by an independent implementation of the same algorithm in C (6.1 s) before the certified run. $\square$

Part (3) is the classical lower bound $R(T_4) \geq 8$ (see [2, 4]), re-proved here from scratch; it is recorded again as Theorem 15. What matters for the ceiling is that (1) and (3) are instances of the *same* predicate, so the strict inequality between them is a fact about one search problem and not a comparison of two techniques.

4.3. The whole family. One member with $R \leq 7$ caps the minimum in $(\star)$ from above. The sweep caps it from below.

Theorem 9. *For every one of the 4096 parameter triples $(\lambda, \mu, \nu) \in (\mathbb{F}_2^4)^3$, the space $V(\lambda, \mu, \nu) \subseteq \mathbb{F}_2^{4 \times 4}$ is contained in the span of no six rank-one matrices. Hence*

$$\min_{\lambda, \mu, \nu \in \mathbb{F}_2^4} R(V(\lambda, \mu, \nu)) = 7,$$

the minimum being attained at $(\lambda, \mu, \nu) = (e_2, e_2, e_2 + e_3)$.

Proof sketch. For each triple the engine is started at an echelon basis of $V(\lambda, \mu, \nu)$ with budget $6 - \dim V(\lambda, \mu, \nu)$ and returns **false**; Theorem 5 converts each **false** into the stated non-existence. This is an exhaustive computation: 103,643,136 candidate subspaces in total across the 4096 triples. The same sweep was first run by an independent C implementation — 0 hits of 4096, 302 s — and the certified run reproduces it. The upper half of the equality is Theorem 8(1); the parameter triple attaining it is a member of the swept family. $\square$

Corollary 10. *One level of the ladder $(\star)$ applied to T_5 yields exactly $R(T_5) \geq 10$, for every choice of the parameters, and never $R(T_5) \geq 11$.*

Proof. Substitute Theorem 9 into $(\star)$ at $N = 4$: the minimum is 7, so the ladder gives $R(T_5) \geq 3 + 7 = 10$ and no more. $\square$

That is exactly the bound of [2, 5] that Wang’s table still prints, and exactly one short of the value that required Covanov’s covering-sets exhaustion. So the ceiling is not an artifact of a weak implementation of the substitution method: the method, in this one-level form, stops where the classical literature stopped. Nor is the witness fragile: Theorem 8(2) shows that the space attaining the minimum admits no enlargement inside the rank-one matrices it contains.

Remark 11 (What is not claimed). Corollary 10 is a statement about one level of $(\star)$ at $N = 4$. It does not say that no substitution-style argument can prove $R(T_5) \geq 11$; it says that this ladder, over its whole parameter family, cannot. It also does not bear on $R(T_5)$ itself, which is 11 by [8].

5. EXACT RANKS BELOW THE CEILING

5.1. Upper bounds. Explicit schemes, each with all N^3 defining identities discharged inside the proof kernel, give the following. All values are known; the schemes were found here (by satisfiability search at $N \leq 6$, and by the span search for the symmetric ones) and are not claimed to be the sources’ own formulae.

Theorem 12. *Over $\mathbb{F}_2$: $R(T_2) \leq 3$, $R(T_3) \leq 5$, $R(T_4) \leq 8$, $R(T_5) \leq 11$, $R(T_6) \leq 14$. By Proposition 3 each bound holds over every commutative $\mathbb{F}_2$ -algebra.*

Theorem 13. *Over $\mathbb{F}_2$, for the cyclic product $C_N = \mathbb{F}_2[x]/(x^N - 1)$: $R(C_3) \leq 4$, $R(C_5) \leq 10$, $R(C_6) \leq 12$, each by a symmetric scheme with all N^3 identities discharged. Soundness (Proposition 2) is proved once for both index conditions, so these bounds too hold over every commutative $\mathbb{F}_2$ -algebra.*

The published record for these cells is $R(C_5) = 10$ and $R(C_6) \in [11, 12]$ [4, 13]; the schemes above certify the upper ends.

The rank-11 scheme for T_5 deserves display, since the number 11 is the value at issue in §4 and none of the sources we consulted prints an explicit $\mathbb{F}_2$ -formula attaining it: [13] attributes the bound to [7] but prints no formula, and [7] itself we could not obtain (no open-access copy exists), so whether a formula appears there is a question we leave open. A symmetric one is: the eleven products are $P_t = L_t(A) L_t(B)$ with the linear forms L_t and the output coefficients they enter given by

$$\begin{array}{llll} L_1 = A_0, & \{0, 1, 2, 3, 4\} & L_7 = A_3, & \{3, 4\} \\ L_2 = A_1, & \{1, 2, 3, 4\} & L_8 = A_0 + A_3, & \{3\} \\ L_3 = A_0 + A_1, & \{1\} & L_9 = A_1 + A_3, & \{4\} \\ L_4 = A_2, & \{2, 3, 4\} & L_{10} = A_4, & \{4\} \\ L_5 = A_0 + A_2, & \{2\} & L_{11} = A_0 + A_4, & \{4\} \\ L_6 = A_1 + A_2, & \{3\} & & \end{array}$$

where the set beside L_t lists the k with $c_{tk} = 1$. All 125 identities are verified in the kernel.

5.2. Lower bounds.

Theorem 14. $R(T_2) = 3$. *The lower bound is certified twice: by a three-identity argument, and by deciding the whole space of rank-2 schemes over $\mathbb{F}_2$ — $(2^4)^3 = 4096$ triples of tables, eight identities each — inside the proof kernel.*

Theorem 15. $R(T_3) = 5$ and $R(T_4) = 8$, both directions certified.

Proof sketch. The upper bounds are Theorem 12. For the lower bounds the general search returns **false** at $(N, r) = (3, 4)$ and at $(N, r) = (4, 7)$, which by Theorem 5 rules out every scheme of those ranks. The searches are exhaustive and small: 49 candidate subspaces at $N = 3$ and 1,867,373 at $N = 4$. Both counts were reproduced in C beforehand. $\square$

The values are classical (§1.1), and their optimality over $\mathbb{F}_2$ is the exhaustive result of [4]. The route matters for what comes next: walking admissible subspaces rather than schemes is what makes the $N = 4$ statement finite in practice, the space of rank-7 schemes at $N = 4$ having 2^{84} elements.

5.3. Negative controls. Exhaustive negative results are worthless if the predicate being exhausted is the wrong one, so each family of statements above is paired with claims that must *fail*, and they are certified to fail.

Proposition 16. *The following are certified false, by evaluation of the check of Lemma 4: the identities are not vacuous, over any nontrivial commutative semiring and any $N \geq 1$ (no rank-0 scheme computes any of these products); the rank-11 scheme for T_5 with its last product deleted does not compute the short product mod x^5 ; reading its recombination table as the right-factor table does not compute it; its tables zero-padded to length 6 do not compute the short product mod x^6 ; the same 0/1 tables fail over $\mathbb{Z}$, so the scheme genuinely uses $1 + 1 = 0$; and the truncated and cyclic tables of the same length do not compute each other's product.*

Proposition 17. *Each search used for a lower bound is paired with a run that must succeed and does: the general search answers **true** at $(N, r) = (3, 5)$ and $(4, 8)$, and the symmetric search answers **true** at every value it must accept, including $(N, r) = (5, 11)$ for the truncated and $(5, 10)$, $(6, 12)$ for the cyclic product. Two degenerate cases are pinned as well: at $r = \dim V$ the search must adjoin nothing and still reject, and at an intermediate rank it must reject. Finally the truncated and cyclic searches are certified to disagree exactly where the two algebras do — symmetric rank 10 is reachable for C_5 and not for T_5 — so they cannot be reading the same tensor.*

The budget-zero control is not decoration: an early prototype of the search descended one level at $r = \dim V$ and reported $R^{\text{sym}}(T_2) \leq 2$ and $R^{\text{sym}}(C_3) \leq 3$, both false. The certified search guards the case, and Proposition 17 is what pins the guard.

6. SYMMETRIC RANKS

The symmetric restriction $b = a$ shrinks the atom set from $(2^N - 1)^2$ rank-one matrices to $2^N - 1$ points uu^T — 63 instead of 3969 at $N = 6$ — and the exhaustive searches shrink with it: at $N = 4$, $r = 7$ the general search walks 1,867,373 candidate subspaces against the symmetric one's 453, and at $N = 5$, $r = 10$ the comparison is $6.8 \cdot 10^{12}$ against 168,693. The subclass is not a curiosity here: the 112 rank-11 formulae that [4] reports for T_5 were found under exactly this restriction — its Table 4 row for $n = 5$ over $\mathbb{F}_2$ carries the mark (Sym.), which [4, §4.6] defines as enumeration over symmetric generators only — so the value at issue in §4 is attained inside the subclass.

Theorem 18. *Over $\mathbb{F}_2$ there are symmetric schemes of ranks 5, 8, 11 and 14 for T_3 , T_4 , T_5 and T_6 , with all N^3 identities discharged in the proof kernel; hence $R^{\text{sym}}(T_3) \leq 5$, $R^{\text{sym}}(T_4) \leq 8$, $R^{\text{sym}}(T_5) \leq 11$, $R^{\text{sym}}(T_6) \leq 14$.*

Theorem 19. $R^{\text{sym}}(T_2) = 3$, $R^{\text{sym}}(T_3) = 5$, $R^{\text{sym}}(T_4) = 8$.

Theorem 20. $R^{\text{sym}}(T_5) = 11$.

Remark 21. Theorem 20 is as close as this machinery comes to the verification question of §1.2. The general value $R(T_5) = 11$ is [8]’s, obtained by the covering-sets method and never independently verified; the 112 rank-11 formulae reported for this cell in [4] were found inside the symmetric subclass, and inside that subclass the value is now certified in both directions, by an explicit scheme and by an exhaustion of 168,693 candidate subspaces. The general lower bound $R(T_5) \geq 11$ remains unverified: Corollary 10 says the cheap route cannot supply it, and Remark 24 prices the expensive one.

We state the scope of that claim exactly, because a stronger one is easy to reach for and we did not check it. What we verified is that the 112 formulae of [4] come from a (Sym.) row, hence from a search restricted to symmetric atoms. The 146,944 optimal decompositions enumerated in [8] were *not* examined for symmetry here, so nothing in this paper should be read as saying that every known optimal T_5 scheme is symmetric.

Theorem 22. $R^{\text{sym}}(C_3) = 4$, $R^{\text{sym}}(C_5) = 10$ and $R^{\text{sym}}(C_6) = 12$.

Theorem 23. $R^{\text{sym}}(T_6) \geq 13$; with Theorem 18, $R^{\text{sym}}(T_6) \in \{13, 14\}$.

Proof sketch. Each upper bound is an explicit symmetric scheme whose N^3 identities are discharged in the kernel. Each lower bound is one exhaustive run of the symmetric search, converted by Theorem 5: an answer of *false* at rank r rules out every symmetric scheme of that rank, and padding (Proposition 3) extends this to all smaller ranks. The searches and their sizes, each reproduced in C before the certified run:

search	answer	candidate subspaces
$T_2, r = 2$	false	1
$T_3, r = 4$	false	7
$T_4, r = 7$	false	453
$T_5, r = 10$	false	168,693
$T_6, r = 12$	false	67,831,173
$C_3, r = 3$	false	1
$C_5, r = 9$	false	31,901
$C_6, r = 11$	false	6,925,299

□

6.1. Provenance of the values. All the values in Theorems 18–23 are published. Table 4 of [4] carries, over $\mathbb{F}_2$, the general rows $\mathbb{F}_2[X]/(X^n)$ with $n = 2, 3, 4$ at $k = 3, 5, 8$ and the (Sym.) rows $n = 5, k = 11$ (112 solutions) and $n = 6, k = 14$ (384 solutions), and for $\mathbb{F}_2[X]/(X^n - 1)$ the general row $n = 3, k = 4$ and the (Sym.) rows $n = 5, k = 10$ (25 solutions) and $n = 6, k = 12$ (31 solutions). The stated convention of that paper is that the reported k is the smallest for which solutions were found, and that in every case “there are no solutions for smaller values of k ”. Four of the values above — $R^{\text{sym}}(T_2) = 3$, $R^{\text{sym}}(T_3) = 5$,

$R^{\text{sym}}(T_4) = 8$, $R^{\text{sym}}(T_5) = 11$ — and the upper end 14 of the T_6 bracket appear again in the $m = 1$ column of Table 2.1 of Rambaud’s thesis [12], which reads $\mu_2^{\text{sym}}(1, \ell) = 1, 3, 5, 8, 11, 14$ for $\ell \leq 6$; that table displays a single entry exactly when the upper bound is optimal, so those are exact values there, and its Table 3.1 credits the upper bounds at $\ell = 5, 6$ to Oseledets [10] and the lower bounds to [4]. The same column is republished, as upper bounds, in [3].

What is new is the certificate. None of the sources cited above ships a machine-checkable artifact: the $n = 6$ cyclic row of [4] alone stands on a $2.33 \cdot 10^7$ -test C program from 2012, and its $n = 6$ truncated row on $2.63 \cdot 10^9$ tests. Every value above is now backed by an explicit scheme whose identities are verified by a proof kernel and, for the lower half, by an exhaustive search whose correctness is itself a proved theorem (Theorem 5).

The word “new” there should be read narrowly. We claim only that no machine-checkable certificate for these values turned up in our own search of the literature — [4] read in full, [8] and [13] read from their e-print sources, [12] and [3] read for their tables, together with keyword sweeps of an arXiv full-text corpus and of the OEIS. That is a statement about what we found, not a proof that no such certificate exists; a reader who knows of one should treat the priority claim as withdrawn and the certificates below as an independent second check, which is worth having in either case.

6.2. Consequences for the two open cells. Wang [13] brackets $R(T_6) \in [13, 14]$, with 13 that paper’s own new lower bound and 14 the upper bound of [7], and $R(C_6) \in [11, 12]$. Theorem 22 says that a rank-11 scheme for C_6 , if one exists, must be non-symmetric: the symmetric value is exactly 12. Theorem 23 is the corresponding half-statement for T_6 : no symmetric scheme of rank at most 12 exists, and [4]’s (Sym.) row at $n = 6$, $k = 14$ says that a rank-13 scheme, if one exists, is non-symmetric as well. Both remarks locate where a closing scheme cannot live, which is what a search should be told before it starts.

7. WHAT THIS DOES NOT REACH

Remark 24 (Measured costs; outside the formal development). The following are measurements of our own tooling, not theorems, and no statement of this paper depends on them.

- An independent verification of $R(T_5) \geq 11$ by the engine of §3 would enumerate $\binom{961}{5} = 6.8 \cdot 10^{12}$ candidate subspaces, about $2 \cdot 10^4$ core-years at the measured node rate. Covanov’s covering-sets method settled the same cell with $6.3 \cdot 10^6$ tests in $2.4 \cdot 10^3$ s — six orders of magnitude fewer, on the reading that the runs at $r < 11$ are the ones its §7.1 calls negligible and does not tabulate. This is the sense in which the missing ingredient is the algorithm rather than the compute, and Corollary 10 is why the cheap substitution route does not supply it.
- $R(C_5) \geq 10$ by the same engine is $\binom{961}{4} = 3.5 \cdot 10^{10}$ subspaces, about 31 CPU-hours; the value is already [4]’s, from $1.39 \cdot 10^{10}$ of their tests.
- $R^{\text{sym}}(T_6) \geq 14$, which would upgrade Theorem 23 to $R^{\text{sym}}(T_6) = 14$ and independently verify [4]’s $2.63 \cdot 10^9$ -test row, is 551,244,164 candidate subspaces — an extrapolated 1.5–2.5 CPU-hours under compiled evaluation. It is affordable and parked, needing only a sharding lemma for the search.

- On the find direction we budgeted about 1.6 CPU-hours across the three open cells — a flip-graph search for T_7 at rank 17, for C_6 at rank 11 and for T_6 at rank 13, and the span search on symmetric T_7 at ranks 16 and 17 — and every run was negative. Controls first: from the schoolbook decomposition the same flip-graph search reaches 11 for T_5 , 14 for T_6 and 18 for T_7 within seconds, so it reproduces every best known value here before failing to beat any of them. The evidence for $R(T_6) = 14$ and $R(C_6) = 12$ keeps accumulating without becoming a proof.

8. VERIFICATION

Every theorem, proposition and lemma of this paper except Proposition 6 has been formalised and machine-checked in Lean 4 with Mathlib. Definitions (schemes, identities, rank bounds, the two index conditions, admissible subspaces) are written once and shared: soundness (Proposition 2) is proved for an arbitrary commutative semiring and mentions no particular scheme, so every per-scheme finite check transfers to every commutative $\mathbb{F}_2$ -algebra along Proposition 3 at no further cost. Upper bounds are discharged by kernel evaluation of the N^3 defining identities — 125 at $N = 5$, 216 at $N = 6$ — and are kernel-clean, using no axioms beyond propositional extensionality, choice and quotient soundness. Lower bounds are the exhaustive searches; each is one compiled evaluation and carries exactly one additional axiom, the one asserting that compiled evaluation agrees with the kernel. Across the development, 58 declarations were audited: none uses `sorry`, 37 are kernel-clean and 21 carry exactly the compiled-evaluation axiom of the search they rest on. The split is the intended one — every upper bound is kernel-only, every lower bound costs one evaluation axiom, and nothing else.

The searches are large enough that the evaluator matters. The engine core imports nothing, so it compiles to a shared library that is loaded into the elaborator; the 67,831,173-subspace run of Theorem 23 then takes 1094.6 s against 174.8 s for the reference C implementation of the same algorithm — about $5.7\times$ C net of the fixed import cost, the expected compiled-Lean factor, where the bytecode interpreter would have been roughly $35\times$ and pushed the same check past 100 minutes. The 103,643,136-subspace sweep of Theorem 9 took 2548.6 s against 302 s in C. Compiled loading is therefore a requirement here, not an optimisation.

Every count quoted in this paper is the node count of an independent C implementation of the same algorithm, written from the reformulation rather than ported; each certified run was made only after that implementation had returned the same answer on the same instance. Its convention is to count the unadjoined seed $W = V$ as one tested subspace. The repository is private: repository reference to be added at submission.

REFERENCES

- [1] A. Alder, V. Strassen, *On the algorithmic complexity of associative algebras*, Theoret. Comput. Sci. **15** (1981), no. 2, 201–211. Not consulted directly; cited through Table 6 of [13].
- [2] A. Averbuch, Z. Galil, S. Winograd, *Classification of all the minimal bilinear algorithms for computing the coefficients of the product of two polynomials modulo a polynomial*, Part I: the algebra $G[u]/\langle Q(u)^\ell \rangle$, $\ell > 1$, Theoret. Comput. Sci. **58** (1988), no. 1–3, 17–56; Part II: the algebra $G[u]/\langle u^n \rangle$, Theoret. Comput. Sci. **86** (1991), no. 2, 143–203. Not consulted directly; cited through the tables of [13].

- [3] S. Ballet, J. Chaumine, J. Pielant, M. Rambaud, H. Randriambololona, R. Rolland, *On the tensor rank of multiplication in finite extensions of finite fields and related issues in algebraic geometry*, arXiv:1906.07456.
- [4] R. Barbulescu, J. Detrey, N. Estibals, P. Zimmermann, *Finding Optimal Formulae for Bilinear Maps*, in: Arithmetic of Finite Fields (WAIFI 2012), Lecture Notes in Computer Science **7369**, Springer, 2012, 168–186; read here in the version posted as IACR Cryptology ePrint Archive, Report 2012/110, whose section and table numbering is the one cited above.
- [5] M. Bläser, *A complete characterization of the algebras of minimal bilinear complexity*, SIAM J. Comput. **34** (2005), no. 2, 277–298. Not consulted directly; cited through the tables of [13].
- [6] P. Bürgisser, M. Clausen, M. A. Shokrollahi, *Algebraic Complexity Theory*, Grundlehren der mathematischen Wissenschaften **315**, Springer, 1997. Ch. 6 is *The Substitution Method*; Ch. 14 is *Multiplicative and Bilinear Complexity*.
- [7] M. Cenk, F. Özbudak, *Multiplication of polynomials modulo x^n* , Theoret. Comput. Sci. **412** (2011), no. 29, 3451–3462. Not consulted directly — no open-access copy exists — and cited only through Table 6 of [13], which attributes to it the upper bounds 11 at $N = 5$ and 14 at $N = 6$ over $\mathbb{F}_2$.
- [8] S. Covanov, *Improved method for finding optimal formulae for bilinear maps in a finite field*, arXiv:1705.07728 (v3, 2018); published as *Improved method for finding optimal formulas for bilinear maps in a finite field*, Theoret. Comput. Sci. **790** (2019) 41–65. Read here from the v3 e-print source. Note that $\text{rank}(\text{ShortProduct}_5) = 11$ appears there in a summary table rather than as a numbered theorem, and that no row records an $r = 10$ run; the reading used in this paper follows that paper’s stated global strategy (§7.1) of increasing r .
- [9] J.-G. Dumas, S. Lia, J. Sheekey, *Computational Explorations on Semifields*, arXiv:2602.09577 (v1, 10 February 2026); the terminology sentence quoted above is in its discussion of the algorithm of [4].
- [10] I. Oseledets, *Optimal Karatsuba-like formulae for certain bilinear forms in $GF(2)$* , Linear Algebra Appl. **429** (2008), no. 8–9, 2052–2066. Not consulted directly; it is reference [18] of [4] and the Oseledets credit in Table 3.1 of [12].
- [11] V. Ya. Pan, *Methods of computing values of polynomials*, Russian Math. Surveys **21** (1966), no. 1, 105–136. Not consulted directly; cited for the origin of the substitution method.
- [12] M. Rambaud, *Courbes de Shimura et algorithmes bilinéaires de multiplication dans les corps finis*, Ph.D. thesis, Télécom ParisTech, 2017 (written in English); Tables 2.1 and 3.1.
- [13] C. Wang, *Automated Lower Bounds for Bilinear Complexity over Finite Fields*, preprint, arXiv:2603.07280 (v10, 30 July 2026). Table numbering is that version’s, obtained by compiling the v10 e-print source: Table 5 is the cyclic convolution modulo $x^N - 1$ and Table 6 the truncated product modulo x^N . In the $\mathbb{F}_2$ columns of Table 6, row $N = 5$ carries previous lower bound 10 (cited there to Averbuch–Galil–Winograd and to Bläser), that paper’s own lower bound 10, and previous upper bound 11 (cited there to Cenk–Özbudak).
- [14] S. Winograd, *Some bilinear forms whose multiplicative complexity depends on the field of constants*, Math. Systems Theory **10** (1977) 169–180. Not consulted directly; cited through Table 6 of [13].