COMS E6998-15: Communication Complexity and Applications (Spring'26) April 22

Matrix Multiplication and NOF Communication

Instructor: *Toniann Pitassi* Presenter: Zachary Thayer

1 Introduction

These notes accompany a presentation of the paper *Matrix Multiplication and Number On the Forehead Communication* by Alman and Błasiok (2023) [1]. The paper observes a striking formal connection: the promise tensor associated with the three-player Number On the Forehead (NOF) communication model is *exactly* the matrix multiplication tensor. This identification lets us transfer ideas developed for the design of fast matrix multiplication algorithms into the setting of communication complexity, and conversely.

Our main goal is to present Theorem 8, which uses Strassen's Laser method [5] to give a non-trivial deterministic protocol for the product of any three NOF permutation problems. Along the way we develop the tensor formalism for NOF and prove the matrix-multiplication identification.

2 Number on the Forehead and Permutation Problems

2.1 The NOF model

In the (three-player) Number on the Forehead model, three players Alice, Bob, and Charlie wish to jointly compute a Boolean function $f : \Sigma^3 \to \{0,1\}$ on input $(x, y, z) \in \Sigma^3$. Each input is written on one player's forehead, so that

$$\text{Alice sees } (y,z), \quad \text{Bob sees } (x,z), \quad \text{Charlie sees } (x,y).$$

Each player sees *everyone's input except their own.* The players communicate over a shared blackboard and want to determine $f(x, y, z)$ using as few bits as possible.

Throughout we take $\Sigma = \{0,1\}^n$, so $N := |\Sigma| = 2^n$. We write $\mathrm{CC}(f)$ for the deterministic communication complexity of f.

2.2 NOF as a promise number-in-hand problem

It will be convenient to recast NOF as a *promise* problem in the Number In Hand (NIH) model. Let each player hold a pair in Σ^2:

$$\text{Alice holds } (x_A, y_A), \quad \text{Bob holds } (y_B, z_B), \quad \text{Charlie holds } (z_C, x_C).$$

The pairs are *promised* to be consistent with a single NOF input, that is,

$$y_A = y_B, \quad z_B = z_C, \quad x_A = x_C.$$

The set of consistent triples of pairs, i.e. tuples of the form $((x,y),(y,z),(z,x))$ for $(x,y,z) \in \Sigma^3$, is called the *NOF promise*, denoted $P_{\text{nof}} \subset \Sigma^2 \times \Sigma^2 \times \Sigma^2$. Any NOF problem can be solved as a promise NIH

problem on $\Sigma^2 \times \Sigma^2 \times \Sigma^2$ with promise P_{nof}, and vice versa.

2.3 Permutation problems

Definition 1 (Permutation problem)**.** *A NOF problem f with accepting set $I = f^{-1}(1) \subset \Sigma^3$ is a* permutation problem *if for each pair $(x, y) \in \Sigma^2$ there is a* unique *$z \in \Sigma$ with $(x, y, z) \in I$, and the analogous statement holds for the other two pairs of coordinates.*

The canonical example from the paper is Eval_G for an abelian group G: $I = \{(x, y, z) \in G^3 : x + y + z = 0_G\}$. For each (x, y), the unique accepting z is $-(x + y)$, and similarly for the other coordinates. In particular, $\text{Eval}_{\mathbb{Z}_N}$ is the famous *corners problem* from additive combinatorics.

Permutation problems all admit a randomized protocol with $O(1)$ communication (each player can guess the unique completion of the other two coordinates and use a randomized equality protocol to check) [1]. They are the natural class of NOF problems for which we hope to prove $\Omega(n)$ *deterministic* lower bounds, and are the focus of this paper.

2.4 Tensor representation

We represent an NOF problem (with promise P_{nof}) as a $\{0, 1, *\}$-valued order-3 tensor T on $\Sigma^2 \times \Sigma^2 \times \Sigma^2$. Index the three axes by Alice's pair, Bob's pair, and Charlie's pair, and set

$$T_{(x_A,y_A),(y_B,z_B),(z_C,x_C)} = \begin{cases} 1 & \text{if the pairs are consistent and } (x, y, z) \in I, \\ 0 & \text{if the pairs are consistent and } (x, y, z) \notin I, \\ * & \text{if the pairs are inconsistent (promise violated).} \end{cases}$$

Even though the ambient tensor has $N^2 \times N^2 \times N^2$ entries, the non-$*$ entries (the valid inputs) live inside an $N \times N \times N$ subtensor, parametrized by the underlying $(x, y, z) \in \Sigma^3$.

Definition 2 (Independent set)**.** *An* independent set *of T is a sub-tensor (obtained by restricting each axis to a subset of indices) all of whose entries are either* 1 *or* $*$*. Equivalently, it is an "all-1's subtensor" once we ignore $*$ entries. We write $I(T)$ for the size of the largest such subtensor (its dimension along any axis, since the subtensor is "diagonal" in the natural sense).*

For a permutation problem, fixing any two of (x, y, z) determines the third, so the structure of independent sets is particularly clean: a subtensor is independent if and only if it is "diagonal," meaning no two of its 1's share any coordinate of (x, y, z).

2.5 Communication complexity from independent sets

The following fact is the bridge between independent sets and communication complexity. We will use it both to upper-bound CC in our main theorem and as the conceptual link to the matrix multiplication side of the story [1].

Fact 3. *For an NOF permutation problem P with promise tensor T,*

$$\text{CC}(P) \leq \log\left(\frac{\#1\text{'s in } T}{I(T)}\right)$$

3 The Matrix Multiplication Tensor

We turn briefly to the world of matrix multiplication, define the matrix multiplication tensor, and recall the connection between its rank and the matrix multiplication exponent ω. We will then observe that the NOF promise tensor we just defined *is* the matrix multiplication tensor.

3.1 Matrix multiplication, formally

Recall standard matrix multiplication: given matrices $A, B \in F^{n\times n}$, their product $C = AB$ has entries

$$C_{i,k} \;=\; \sum_{j=1}^{n} A_{i,j} \cdot B_{j,k}.$$

Definition 4 (Matrix multiplication tensor)**.** *The $n \times n$* matrix multiplication tensor *$\langle n, n, n\rangle$ is the order-3 $\{0,1\}$-valued tensor on $F^{n^2} \otimes F^{n^2} \otimes F^{n^2}$ given by*

$$\langle n, n, n\rangle \;=\; \sum_{i,j,k=1}^{n} x_{i,j}\, y_{j,k}\, z_{k,i},$$

where $x_{i,j}, y_{j,k}, z_{k,i}$ stand for the entries of A, B, C respectively.

Equivalently, indexing the three axes by pairs $(i,j),(j',k),(k',i')$, we have

$$\langle n, n, n\rangle_{\,(i,j),(j',k),(k',i')} \;=\; \begin{cases} 1 & \text{if } i = i',\ j = j',\ k = k', \\ 0 & \text{otherwise.} \end{cases}$$

More generally, $\langle a, b, c\rangle$ denotes the tensor for multiplying an $a \times b$ matrix by a $b \times c$ matrix.

A useful identity for products is

$$\langle a_1, b_1, c_1\rangle \otimes \langle a_2, b_2, c_2\rangle \;=\; \langle a_1 a_2, b_1 b_2, c_1 c_2\rangle, \tag{1}$$

which we will use repeatedly.

3.2 Fast matrix multiplication and the exponent ω

Definition 5 (Tensor rank)**.** *A tensor is* rank one *if it has the form $u \otimes v \otimes w$ for some vectors u, v, w. The* rank *$R(T)$ of an order-3 tensor T is the smallest r such that T is a sum of r rank-one tensors.*

A rank-r decomposition of $\langle n, n, n\rangle$ corresponds to a bilinear algorithm computing $n \times n$ matrix multiplication using r scalar multiplications [4]. The *matrix multiplication exponent* is

$$\omega \;:=\; \inf\left\{ t \;:\; R(\langle n, n, n\rangle) = O(n^t) \right\}.$$

Trivially $\omega \le 3$ from the schoolbook algorithm, and $\omega \ge 2$ since the output has n^2 entries. Any rank-r decomposition of $\langle q, q, q\rangle$ for fixed q implies $\omega \le \log_q r$.

Example 6 (Strassen, 1969)**.** *Strassen exhibited a decomposition of $\langle 2, 2, 2\rangle$ into a sum of 7 rank-one tensors, beating the trivial $r = 8$ from the schoolbook algorithm [4]. This gives $\omega \le \log_2 7 \approx 2.81$, and*

was the first nontrivial upper bound on ω. Subsequent improvements (Coppersmith–Winograd and many others) have brought the bound down towards 2, all using more sophisticated algebraic techniques, the Laser method that we will encounter in the proof of Theorem 8 is one of the core such techniques [3, 5].

3.3 The key observation: $P_{\text{nof}} = \langle N, N, N \rangle$

We now arrive at the conceptual punchline of the paper [1]. Recall the NOF promise tensor P_{nof} from §2. The accepting entries are precisely the tuples $((x,y),(y,z),(z,x))$ for $(x,y,z) \in \Sigma^3$; that is, P_{nof} has a 1 at index $((a_1,b_1),(a_2,b_2),(a_3,b_3))$ iff $b_1 = a_2$, $b_2 = a_3$, $b_3 = a_1$.

Comparing with the explicit formula for the matrix multiplication tensor $\langle N,N,N\rangle_{(i,j),(j',k),(k',i')} = \mathbf{1}[i = i', j = j', k = k']$, we see that the constraints match *exactly*, with the role of (i,j,k) played by (x,y,z).

Observation 7. $P_{\text{nof}} = \langle N, N, N \rangle$.

This formal identification is the soul of the paper: from now on, anything we know about the structure of the matrix multiplication tensor (its rank, subrank, independent sets, etc.) we can deploy in the analysis of NOF problems.

4 A Non-Trivial Protocol via the Laser Method

We now prove the main result. Recall that the trivial bound for solving any NOF problem on n-bit inputs is n bits of communication (one player just sends their entire pair). For *three* independent NOF problems with n-bit inputs, the trivial bound is therefore $3n$. The following theorem says we can do dramatically better when we want to solve all three simultaneously (compute their AND).

Theorem 8 (Alman–Błasiok [1])**.** *Let P, Q, R be three NOF permutation problems, each on alphabet $\Sigma = \{0,1\}^n$. Then*

$$\mathrm{CC}(P \otimes Q \otimes R) \;\le\; (1 + o(1))\, n.$$

The trivial bound of $3n$ is beaten by an asymptotic factor of 3.

4.1 Strategy: build one large independent set

By Fact 3, an upper bound on $\mathrm{CC}(P \otimes Q \otimes R)$ follows from a lower bound on I of its promise tensor. Each of P, Q, R is a permutation problem on alphabet of size N, so each has N^2 accepting entries; under Kronecker product these multiply, giving N^6 accepting entries in the tensor for $P \otimes Q \otimes R$.

Our goal is to show $I(P \otimes Q \otimes R) \ge N^{5-o(1)}$, from which Fact 3 gives

$$\mathrm{CC}(P \otimes Q \otimes R) \;\le\; \log\!\left(\frac{N^6}{N^{5-o(1)}}\right) + o(\log N) \;=\; \log\!\left(N^{1+o(1)}\right) \;=\; (1+o(1))n.$$

4.2 Block tensors: outer and inner structure

The Laser method analyzes a tensor by partitioning each of its axes into blocks and treating the resulting "macro" structure (outer) and "micro" structure (inner) separately.

Definition 9 (Block decomposition)**.** *Let T be a tensor with axes A, B, C. A* block decomposition *of T consists of partitions $A = \bigsqcup_i A_i$, $B = \bigsqcup_j B_j$, $C = \bigsqcup_k C_k$. We define:*

- *the* outer tensor $\mathrm{Out}(T)$*, indexed by (i,j,k), with $\mathrm{Out}(T)_{i,j,k} = 1$ iff the block $T|_{A_i \times B_j \times C_k}$ is nonzero, and 0 otherwise;*
- *the* inner tensors*, namely $\mathrm{In}(T)_{i,j,k} := T|_{A_i \times B_j \times C_k}$ for each nonempty block.*

The following is the basic combinatorial fact that makes the Laser method possible [5, 1]: independent sets of the whole tensor can be lower-bounded by combining an independent set in the outer tensor with independent sets in the inner tensors.

Lemma 10 (Outer-inner lemma)**.** *For any block decomposition of a tensor T,*

$$I(T) \;\geq\; I(\mathrm{Out}(T)) \cdot \min_{T' \in \mathrm{In}(T)} I(T').$$

4.3 Choosing the partition

We now describe a specific block decomposition for each of P, Q, R. Recall that as tensors, the accepting entries of P have the form $((i,j),(j,k),(k,i))$ where (i,j,k) ranges over the accepting set of the underlying NOF problem.

Partition for P: fix the j-coordinate. On the first axis (indexed by pairs (i,j)) we partition by j:

$$A_t \;:=\; \{(i,t) : i \in [N]\} \quad \text{for each } t \in [N].$$

On the second axis (indexed by (j,k)), we partition by j:

$$B_t \;:=\; \{(t,k) : k \in [N]\} \quad \text{for each } t \in [N].$$

The third axis (indexed by (k,i)) is left as a single block.

We can visualize this as cutting the cube $[N^2] \times [N^2] \times [N^2]$ into "slabs" along the j-direction: the t-th pair of slabs $A_t \times B_t \times C$ is exactly the part of the cube where the j-coordinate is fixed to t.

Partitions for Q and R. By the same construction (relabeling axes), we choose for Q a partition that fixes k, and for R a partition that fixes i.

4.4 Outer structure

We analyze the outer structure of the partition for P; the analyses for Q and R are completely symmetric.

A block $A_{t_1} \times B_{t_2} \times C$ contains an accepting tuple $((i,j),(j,k),(k,i))$ iff $j = t_1$ (forced by A_{t_1}) *and* $j = t_2$ (forced by B_{t_2}). So the block is nonempty iff $t_1 = t_2$. We get

$$\mathrm{Out}(P) \;=\; \langle N, 1, 1 \rangle.$$

By symmetry,

$$\mathrm{Out}(Q) \;=\; \langle 1, 1, N \rangle, \qquad \mathrm{Out}(R) \;=\; \langle 1, N, 1 \rangle.$$

Now using the product identity (1),

$$\mathrm{Out}(P \otimes Q \otimes R) \;=\; \mathrm{Out}(P) \otimes \mathrm{Out}(Q) \otimes \mathrm{Out}(R) \;=\; \langle N, 1, 1\rangle \otimes \langle 1, 1, N\rangle \otimes \langle 1, N, 1\rangle \;=\; \langle N, N, N\rangle.$$

This is where the matrix multiplication tensor reappears! And we now get to use a famous theorem of Strassen to control its independent set size.

Theorem 11 (Behrend [2])**.** $I(\langle N, N, N\rangle) \;\geq\; N^{2-o(1)}$.

We omit the proof but note that this result was largely influential in the progression of ω.

Combining, $I(\mathrm{Out}(P \otimes Q \otimes R)) \;\geq\; N^{2-o(1)}$.

4.5 Inner structure

Now we turn to the inner tensors. Take any nonempty block of the partition for P, namely $A_t \times B_t \times C$. Its accepting tuples are precisely those of the form

$$((i,t),\,(t,k),\,(k,i)) \quad \text{for } i,k \in [N].$$

As (i,k) ranges over $[N] \times [N]$, this slab has N^2 entries in the ambient tensor; the 1's among them are determined by the underlying NOF problem.

Here is where the *permutation property* of P enters in an essential way. For each i and each t, there is a unique k such that (i,t,k) is accepting. So the block contains exactly N accepting entries (one per i). Moreover, by the permutation property applied across all three coordinate pairs, no two of these N accepting entries share an i, a t, or a k. Hence these N entries themselves form an independent set:

$$I(\mathrm{In}(P)) \;=\; N.$$

The same analysis applies to Q and R. Independent sets multiply under Kronecker product, so

$$I(\mathrm{In}(P \otimes Q \otimes R)) \;\geq\; I(\mathrm{In}(P)) \cdot I(\mathrm{In}(Q)) \cdot I(\mathrm{In}(R)) \;=\; N \cdot N \cdot N \;=\; N^3.$$

4.6 Putting it together

We now combine the outer and inner bounds via Lemma 10:

$$I(P \otimes Q \otimes R) \;\geq\; I(\mathrm{Out}(P \otimes Q \otimes R)) \cdot \min_{T' \in \mathrm{In}(P \otimes Q \otimes R)} I(T') \;\geq\; N^{2-o(1)} \cdot N^3 \;=\; N^{5-o(1)}.$$

Since each of P, Q, R has N^2 accepting entries and these multiply under $\otimes$, the total number of 1's in $P \otimes Q \otimes R$ is N^6. By Fact 3,

$$\mathrm{CC}(P \otimes Q \otimes R) \;\leq\; \log\!\left(\frac{N^6}{N^{5-o(1)}}\right) + o(\log N) \;=\; \log\!\left(N^{1+o(1)}\right) \;=\; (1+o(1))\,n,$$

which proves Theorem 8. □

References

[1] Josh Alman and Jarosław Błasiok. Matrix multiplication and number on the forehead communication. *arXiv preprint arXiv:2302.11476*, 2023.

[2] Felix A. Behrend. On sets of integers which contain no three terms in arithmetical progression. *Proceedings of the National Academy of Sciences*, 32(12):331–332, 1946.

[3] Don Coppersmith and Shmuel Winograd. Matrix multiplication via arithmetic progressions. In *Proceedings of the Nineteenth Annual ACM Symposium on Theory of Computing*, pages 1–6, 1987.

[4] Volker Strassen. Gaussian elimination is not optimal. *Numerische Mathematik*, 13(4):354–356, 1969.

[5] Volker Strassen. Relative bilinear complexity and matrix multiplication. *Journal für die reine und angewandte Mathematik*, 375:406–443, 1987.