CS 6998 Final Project: Matrix Multiplication, Additive Combinatorics, and Number on Forehead Communication

Eric Liu

December 20, 2023

Abstract

This project will be an exposition on some of the connections between Matrix Multiplication, Additive Combinatorics, and Number on Forehead communication. The goal is to give many concrete examples of the reductions done between these three areas, and to explore some ways in which solving a problem in one area can lead to improvements in another.

1 Introduction

The Number on Forehead (NOF) communication model, first introduced by Chandra, Furst, and Lipton [1], is a model in communication complexity where there are $k \geq 3$ players, who are allowed to see the inputs given to all the other players, but not their own. Letting Σ be the domain, the goal is to jointly compute a function $f : \Sigma^k \to \{0, 1\}$ while minimizing communication.

Since its introduction, the NOF model has had an extremely tight coupling with the field of additive combinatorics, which is a field that deals with counting additive structures in groups. For example, two structures relevant to this paper are three-term arithmetic progressions, and *corners*. Although the former is much more well understood now, there is still much work to be done on understanding the latter. Often, a NOF communication problem has a precise correspondence with problems in additive combinatorics, including the two mentioned before, and for this reason progress in NOF communication has been relatively difficult. Conversely, recent work has actually used the NOF perspective in order to improve longstanding bounds on the Corners problem [2] [3].

In the area of matrix multiplication, many connections to additive combinatorics are also known. For example, the cap-set problem was used to show barriers in using the group-theoretic approach to proving the matrix multiplication exponent $\omega = 2$. Conversely, Behrend's construction [6] was used both in Coppersmith-Winograd's laser method to show $\omega \leq 2.387$ [7], and also in the initial NOF paper by Chandra, Furst, and Lipton.

Recently, a connection between three-player NOF communication and matrix multiplication was established by Alman and Błasiok in [4]. This connection, which involves associating the input (a subset of Σ^3) with its three-dimensional tensor, allows us to use well understood algebraic techniques in matrix multiplication to better understand problems in NOF communication.

2 Number on Forehead

For the rest of this paper we will focus on the case where the number of parties $k = 3$. This section will further describe a set of relationships between NOF and additive combinatorics, which were summarized/discovered by the authors in [2], [1], and [3].

One of the central problems in NOF communication is to find a function which is easy given shared randomness, but hard for deterministic protocols. In the two-party case, where NOF is equivalent to

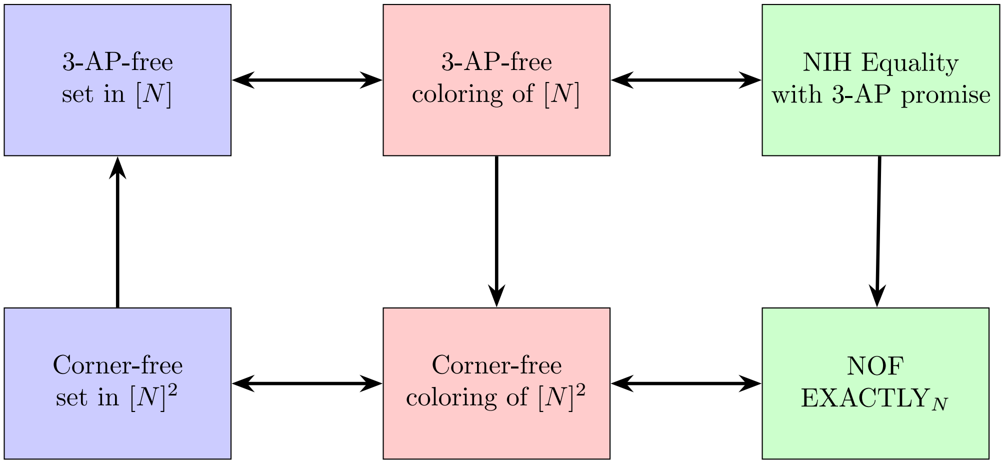
3-AP-free set in $[N]$
3-AP-free coloring of $[N]$
NIH Equality with 3-AP promise
Corner-free set in $[N]^2$
Corner-free coloring of $[N]^2$
NOF EXACTLY$_N$

Figure 1: Relations (taken from [3]). For problems A and B, $A \to B$ denotes $c(B) = O(c(A))$, where $c(\cdot)$ measures the problem's complexity in our context.

the classical model of two-party communication, there exists such a strong separation. If each player is given a number in $[N]$, there is a randomized protocol which takes $\mathcal{O}(1)$ communication, but the deterministic protocol is known to be $\Omega(\log N)$. It is known by a non-constructive argument [8] that there exists a problem which satisfies the same separation, so the goal now is to show an explicit such problem.

An analogous candidate to exhibit such a separation in the three-party NOF case is the EXACTLY$_N$ problem, where the players each have as input (i.e. the number on their forehead) $x, y, z \in [N]$, and are asked to accept if $x + y + z = N$, and otherwise reject. The randomized protocol for this problem uses the randomized protocol for equality in the two-party case. Player 1 sees the values (x, y) on Player 2 and Player 3's foreheads, and simply asks Player 2 if the number on her (Player 1's) own head is equal to $N - (x + y)$. On the other hand, the best known lower bound for this problem is $\Omega(\log \log \log N)$, which is far from what we seek.

Having reached an impasse, we turn our attention to the converse: what are the *upper* bounds on the deterministic protocol and what are the implications? We will now explain the reductions between problems in additive combinatorics and the NOF EXACTLY$_N$ problem.

We start by giving the definitions needed to understand the figure above. A three-term arithmetic progression (3-AP) is defined as a tuple $(x, x + \delta, x + 2\delta) \in [N]^3$ with $\delta \neq 0$. A corner set in $[N]^2$ is defined as a tuple $(x, y), (x + \delta, y), (x, y + \delta)$ with $\delta \neq 0$. The problems in blue thus are asking for the largest subset of $[N]$ (resp. $[N]^2$) without any 3-APs (resp. corners). The coloring versions of these problems ask for the minimum number of colors to partition $[N]$ (resp. $[N]^2$) into such that each color class does not have any 3-APs (resp. corners). A standard probabilistic tiling argument [3] gives that a 3AP-free subset of size N/δ implies the existence of a 3-AP-free coloring using $\delta O(\log N)$ colors. Thus, a lower bound on the 3-AP-free set problem implies an upper bound in the coloring problem (and of course analogous results hold for the Corners problem).

We mentioned earlier that corners are less well understood than arithmetic progressions. One instantiation of this is that we can construct a Corner-free set from a 3-AP-free set. This can be seen by mapping $S \subset [N]$ to $\{(x, y) : x - y \in S\} \subset [N]^2$. A corner $((x, y), (x+\delta, y), (x, y+\delta))$. In particular, we have that the largest size of a Corner-free set in $[9N]$ is at least N times the size of the largest 3-AP-free set in $[N]$. This is modeled in the left vertical arrow in the figure above.

Recall that we wished to show the implications of an upper bound the communication complexity for the NOF EXACTLY$_N$ problem. We will do so by showing the right-to-left implications in the bottom row: an upper bound on the deterministic protocol for NOF EXACTLY$_N$ gives a lower bound for the Corner-free set problem.

Theorem 1 ([1])**.** Let c be the NOF communication complexity of the Exactly$_N$ problem with 3 players, and let $c_\angle(N)$ be the minimum number of colors needed to color $[N]^2$ such that no corner is monochromatic. The following holds:

$$\log c_\angle(N/2) \leq c \leq \log c_\angle(N) + 2$$

We will show the lower bound here:

Proof. Let π be a NOF protocol for EXACTLY$_N$. We will produce a Corner-free coloring of $[N/2]^2$ using π. Color each $(x, y) \in [N]^2$ by the transcript of π on the input $(x, y, N - (x + y))$. Suppose for the sake of contradiction that there were a monochromatic corner $(x, y), (x + \delta, y), (x, y + \delta)$ for $\delta \neq 0$. Then $(x, y, N-(x+y)), (x+\delta, y, N-(x+y+\delta)), (x, y+\delta, N-(x+y+\delta))$ all have the same transcript. But this implies that $(x, y, N - (x + y + \delta))$ also has the same transcript. This is because, for example, Player 3 only sees (x, y), and cannot distinguish between whether she has $N-(x+y)$ or $N-(x+y+\delta)$ on her forehead. But since a similar situation holds for all players, and all transcripts are the same, this means $(x, y, N - (x + y + \delta))$ also has the same transcript. However, this is a contradiction, since $x + y + N - (x + y + \delta) = N - \delta \neq N$, whereas our transcript should be accepting. □

Thus, an upper bound on the NOF communication complexity of EXACTLY$_N$ implies a lower bound on the Corner-free coloring of $[N/2]^2$. Let $r_\angle(N)$ be the maximum size of a Corner-free subset of $[N]^2$. Because each coloring class is a Corner-free set, by the pigeonhole principle we have that $r_\angle(N) \geq N/c_\angle(N)$. We now have a chain of inequalities that shows an upper bound on the deterministic communication complexity of EXACTLY$_N$ implies a lower bound on the Corner-free sets problem.

The original construction of Behrend gives an upper bound of $(2\sqrt{2} + o(1))\sqrt{\log N}$ on the deterministic NOF communication complexity of EXACTLY$_N$ via the reductions in the top row and then the right vertical arrow. That is, we started with a construction of a 3-AP-free set in $[N]$ of size $2^{2\sqrt{2}\sqrt{\log N}+o(\sqrt{\log N})}$, which in turn gave an upper bound on the coloring version of the problem, which we then converted into upper bounds on the communication problems. In fact, for a long time, the best lower bound for Corner-free set was due to this this series of reductions. However, recent work in [2] [3] aimed to improve this by giving explicitly NOF protocols for EXACTLY$_N$. They did this by giving a constructive version of Behrend's construction (which non-constructive since it used the pigeonhole principle), and then further by exploiting shared information.

3 Matrix Multiplication

We now turn our attention to the use of algebraic techniques. The main idea leveraged by Alman and Błasiok [4] is the formal connection between NOF problems and the matrix multiplication tensor. A three-party NOF problem over inputs in Σ^3 can equivalently be viewed as a *Promise Number in Hand* (NIH) problem over $(\Sigma \times \Sigma)^3$. The idea is just that instead of considering the number on the forehead as the input to a player, instead consider what the player sees, i.e. the numbers on the other two foreheads, as the input. More explicitly, map an NOF input (a, b, c) to a Promise NIH input $((b, c), (c, a), (a, b))$. This is intuitive since any protocol that the player performs can only use the other two numbers as input anyways. The only catch is that not every element of $(\Sigma \times \Sigma)^3$ corresponds to a NOF problem. We need to restrict our inputs to be elements of $P = \{((b, c), (c, a), (a, b)) : a, b, c \in \Sigma\}$.

In the NOF setting, we had a subset $I \subseteq \Sigma^3$ of instances such that we wanted to accept if the input was in I and reject otherwise. Using the same mapping as before, have a subset $\tilde{I} \subseteq P \subseteq (\Sigma \times \Sigma)^3$.

The key idea is to not just treat these as subsets of $(\Sigma \times \Sigma)^3$, but actually to consider them as $\{0, 1\}$ order-3 tensors in $\mathbb{F}^{(\Sigma\times\Sigma)^3}$. Thus, we have two tensors T_I and T_P, where T_I is problem-dependent and $T_P = \sum_{a,b,c\in\Sigma} x_{(a,b)}y_{(b,c)}z_{(c,a)}$. The observation is that T_P is exactly the matrix multiplication tensor.

Using this connection, they proved a host of results both about the matrix multiplication tensor and in NOF complexity.

A key definition is that of the *permutation problem.* We say that $\tilde{I} \subset P \subseteq (\Sigma \times \Sigma)^3$ is a permutation problem if the tensor $T_I := \sum_{(i,j,k)\in \tilde{I}} x_i y_j z_k$ restricted to the coordinates appearing in I (i.e. $(T_I)_{\pi_1(I),\pi_2(I),\pi_3(I))}$ is a permutation of an identity tensor. The last condition is that $\tilde{I} = |\Sigma \times \Sigma|$. In words: for every valid input of the $i \in \Sigma$ of the first player, there is at most one pair j, k, such that $(i, j, k) \in \tilde{I}$, and analogously for the second and the third player.

An example of a permutation problem is EVAL_N, (which is just the group version of $\mathrm{EXACTLY}_N$ and has basically the same communication complexity), in which the players are given $x, y, z \in \mathbb{Z}_N$ and asked to accept if and only if $x + y + z = 0$.

Using a counting argument, they are able to show that for most of NOF permutation problems, over the alphabet $[N]$ where $N = 2^n$ there is a lower bound of $\frac{n}{3} - \mathcal{O}(1)$ deterministic communication. Conversely, by the laser method, they are able to show that the *asymptotic* protocol of any NOF permutation problem has an upper bound of $(1+o(1))\frac{n}{3}$ deterministic communication. The lower bound does not require any algebraic technique; it simply uses the setup of being a permutation problem to give a counting argument. On the other hand, the upper bound heavily uses the algebraic structure of the problem: the fact that the promise tensor is the multiplication tensor implies the existence of outer structure and inner structure of T_I, and the fact that we are dealing with a permutation problem makes the inner structure "easy". (The actual proof method is by giving a coloring, where each color is a pair of a color for the outer structure tensor and a color for the inner structure tensor. The permutation problem is defined in such a way that it only requires one color.)

Up until now, we have defined our communication problem with a pair $(\tilde{I}, P)$, where P was the matrix multiplication tensor. However, we could also consider other promises. For example, in Figure 1 above, we had in the top right rectangle the 3-AP promise as a part of the reduction to the $\mathrm{EXACTLY}_N$ protocol.

We will now briefly discuss a rather general, yet extremely powerful class called the asymptotic spectrum. These are a class of functions $r : \mathbb{F}^{\Sigma\times\Sigma\times\Sigma} \to \mathbb{R}_{\geq 0}$ that generalize the notion of matrix rank to tensors. In particular, r is required to satisfy (1) being sub-additive for direct sums, and (2) if B is an identification of A, then $r(B) \leq r(A)$, where an identification is done by zeroing out and then setting some variables equal to each other. We that $\mathcal{CC}(\tilde{I}, P)$, the deterministic communication complexity of the problem $(\tilde{I}, P)$ is equal to the minimum c such that $T_I = \sum_{i=1}^{2^c} T_i$, where each T_i is a zeroing out of P. This is because each protocol of length c induces a partition of 2^c combinatorial rectangular prisms (analogous to the two-party case, where a partition into rectangles is induced). Then we have a theorem that states that

Theorem 2. If r is a part of the asymptotic spectrum, then for any problem three-party Promise NIH problem (I, P), we have

$$\mathcal{CC}(I, P) \geq \log \frac{r(I)}{r(P)}$$

Proof. Let $k = 2^{\mathcal{CC}(I,P)}$. By the above, $T_I = \sum_{i=1}^{k} T_i$, where each T_i is a zeroing out of P. In particular, the direct sum of k copies of P has a zeroing out to the direct sum of all the T_i tensors, which in turn has an identification to I. We thus have

$$r(I) \leq \sum_{i=1}^{k} r(T_i) \leq k \cdot r(P)$$

as desired. □

In particular, tensor rank, slice-rank, and zero-out subrank satisfy the conditions for r. This theorem gives a nice understanding of how exactly I and P relate to each other. In order to prove a lower bound, one would want $r(I) >> r(P)$. One immediate result is that since the matrix multiplication tensor has maximal slice rank, we cannot prove lower bounds this way for NOF communication problems!

Now we turn to the NIH Equality with 3-AP promise, except importantly, we do not work over $[N]$, but instead $\mathbb{Z}_3^N$. In this model, the three players are given $x, y, z \in \mathbb{Z}_3^N$ with the promise that $x + y + z = 0$, and want to determine if $x = y$. One could use the set of reductions in the top row of Figure 1, in order to show that this is equivalent to the 3-AP-free set problem in $\mathbb{Z}_3^N$ in the sense that upper bounds on the size of 3-AP-free sets (also known as Cap sets, when we are working over $\mathbb{Z}_3^N$) imply lower bounds on the deterministic communication complexity of the given problem, and vice versa. Giving a good upper bound on the asymptotic growth of Cap sets was a longstanding conjecture which attracted the attention of many famous mathematicians, and was famously solved in the last decade by using an extremely clever application of the polynomial method [9] [10].

An "alternative" way of showing the equivalence above is by instead using the above theorem. In particular, the Cap Set Theorem tells us that $r(P) \leq 2.756^N$, but by virtue of being a permutation problem, $r(I) = 3^N$ (where r is the zero-out subrank). Hence, we have an $\Omega(N)$ lower bound. Interestingly enough, there is no known direct proof of this lower bound.

4 Group-Theoretic Approach to Matrix Multiplication

In this final section, we will briefly overview how results in additive combinatorics connect to matrix multiplication. As a connection to what was previously discussed here, we recall that Behrend's construction, which was used to give an efficient deterministic protocol for the EXACTLY$_N$ problem, was also used in the laser method for matrix multiplication.

We will now turn our attention to another approach, which is the group-theoretic approach. The idea is that given a group and some triplets of subsets in the group, if the subsets satisfy the so-called simultaneous triple product property (STPP), then we can reduce several independent matrix multiplications into a single group algebra multiplication [11]. In the work of Blasiak et al. [12], they showed a barrier to using this method for groups of bounded exponent by showing a contradiction. They proved this by showing (1) any abelian group H satisfying the STPP must have a multicolored sum-free set of a certain size, and (2) proving an upper bound on the size of multi-colored sum-free sets in $\mathbb{F}_p^n$ (for fixed p), which contradicted the lower bound. Their proof of (2) was done by proving a tri-colored version of the work of Ellenberg-Gijswijt.

Recently, a preprint by Pratt [13], the author proposed several conjectures in additive combinatorics which would provide more barriers to the group-theoretic approach. In particular, they would rule out using groups with a bounded number of direct factors, which is in contrast to the previous barrier. The motivating observation was that the matrix multiplication hypergraph is an extremal solution to some forbidden hypergraph problem. Due to this observation, they were able to derive a whole host of shape-related conditions which somehow relate to matrix multiplication.

The first such proposal made is about the largest size of a set without *skew* corners: a set of three points of the form $(x, y), (x, y + \delta), (x + \delta, y')$. Given our discussion earlier in the paper, would it be possible to design a three-party communication problem which captures this? The difficulty here is that one of the players must compute all possible colors that match their input along a line, which seems to contradict the purpose of coloring.

5 Conclusion

In this paper, we discussed many reductions, and showed how viewing the same problem through a different lens can often give breakthroughs. Recently, a lot of work has been done on either improving

NOF protocols, or on giving lower bounds. To computer scientists, thinking in this type of setting may be the most natural, and hence we want to highlight this as an exciting direction for future work.

References

[1] Ashok K Chandra, Merrick L Furst, and Richard J Lipton. Multiparty protocols. In *Proceedings of the fifteenth annual ACM symposium on Theory of computing*, pages 94–99, 1983.

[2] Nati Linial, Toniann Pitassi, and Adi Shraibman. *On The Communication Complexity of High-Dimensional Permutations.* 2018. `https://arxiv.org/abs/1706.02207`. *arXiv preprint*, cs.CC.

[3] Lianna Hambardzumyan, Toniann Pitassi, Suhail Sherif, Morgan Shirley, and Adi Shraibman. *An improved protocol for ExactlyN with more than 3 players.* 2023. `https://arxiv.org/abs/2309.06554`. *arXiv preprint*, cs.CC.

[4] Josh Alman and Jarosław Błasiok. *Matrix Multiplication and Number On the Forehead Communication.* 2023. `https://arxiv.org/abs/2302.11476`. *arXiv preprint*, cs.CC.

[5] Jonah Blasiak, Thomas Church, Henry Cohn, Joshua A. Grochow, Eric Naslund, William F. Sawin, and Chris Umans. *On cap sets and the group-theoretic approach to matrix multiplication. Discrete Analysis*, Alliance of Diamond Open Access Journals, Jan 2017. `http://dx.doi.org/10.19086/da.1245`. DOI: 10.19086/da.1245. ISSN: 2397-3129.

[6] F. 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.

[7] 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.

[8] Paul Beame, Matei David, Toniann Pitassi, and Philipp Woelfel. *Separating deterministic from randomized multiparty communication complexity. Theory of Computing*, 6(1):201–225, 2010.

[9] Jordan S. Ellenberg and Dion Gijswijt. *On large subsets of F_q^n with no three-term arithmetic progression.* 2016. `https://arxiv.org/abs/1605.09223`. *arXiv preprint*, math.CO.

[10] Ernie Croot, Vsevolod Lev, and Peter Pach. *Progression-free sets in $\mathbb{Z}_4^n$ are exponentially small.* 2016. `https://arxiv.org/abs/1605.01506`. *arXiv preprint*, math.NT.

[11] H. Cohn, R. Kleinberg, B. Szegedy, and C. Umans. *Group-theoretic Algorithms for Matrix Multiplication.* In *46th Annual IEEE Symposium on Foundations of Computer Science (FOCS'05)*, IEEE, 2005. `http://dx.doi.org/10.1109/SFCS.2005.39`. DOI: 10.1109/SFCS.2005.39.

[12] Jonah Blasiak, Thomas Church, Henry Cohn, Joshua A. Grochow, Eric Naslund, William F. Sawin, and Chris Umans. *On cap sets and the group-theoretic approach to matrix multiplication. Discrete Analysis*, Alliance of Diamond Open Access Journals, Jan 2017. `http://dx.doi.org/10.19086/da.1245`. DOI: 10.19086/da.1245. ISSN: 2397-3129.

[13] Kevin Pratt. *On generalized corners and matrix multiplication.* 2023. `https://arxiv.org/abs/2309.03878`. *arXiv preprint*, math.CO.