Probabilistic and experimental method in Sum-Product Theory

Xiaobo Luo

April 19, 2021

1 Introduction

Let's start with some necessary definitions

Definition 1 (Sumset). *Let $A, B \subseteq G$ where $(G, +)$ is an Abelian Group, then the sumset of A, B is defined to be $A + B = \{a + b \mid a \in A, b \in B\}$.*

Definition 2 (Product Set). *Let $A, B \subseteq G$ where $(G, \cdot)$ is an Abelian Monoid, then the product set of A, B is defined to be $A \cdot B = \{a \cdot b \mid a \in A, b \in B\}$.*

Both definitions are defined in general setting and we are normally working with the construction of G being $\mathbb{Z}$ or $\mathbb{Z}_N$ (Multiplicative group of integers modulo n) or $\mathbb{R}$.

The sum set and product set is first investigated by Erdos and Szemeredi[1] in 1983. In their paper, they proved that for $A \subseteq \mathbb{Z}$ being a set of integers, then

$$\max(|A + A|, |A \cdot A|) \geq c|A|^{1+\epsilon}$$

for some small and positive ϵ where $|\cdot|$ denotes the size of the set. They further conjectured that it should be the case

$$\max(|A + A|, |A \cdot A|) \geq c|A|^{2-\epsilon}$$

for any positive ϵ.

This problem is further analyzed in the setting where $A \subseteq \mathbb{R}$ and after works by Nathanson[2], Ford[3] and Chang[4], Elekes[5] shows that $\epsilon \geq \frac{1}{4}$. This result is further extended to complex numbers by Toth[6] and Solymosi[7]. The best know bound is proven by Solymosi[8] which is

$$\max(|A + A|, |A \cdot A|) \geq \frac{c|A|^{\frac{14}{11}}}{\log^{\frac{3}{11}} |A|}$$

This is further analyzed in the setting of finite field but the situation becomes more complex as the key tool used in the analysis, Szemeredi-Trotter incidence theorem, doesn't hold in the same generality. It is rst shown by Bourgain, Glibichuk and Konyagin [9] and Bourgain, Katz and Tao[10] that if q is a prime, than if $|A| \le Cq^{1-}$, for some > 0, then there exists > 0 such that

$$\max\{|A + A|; |A\,A|\} \ge c|A|^{1+}$$

Hart, Iosevich and Solymosi [11] further improved this bound using incidence theorem and get

$$\max\{|A + A|; |A\,A|\} \ge c|A|^{\frac{8}{7}}$$

This gives us a better understanding of the size of $|A + A|$ and $|A\,A|$. In this paper, we will look at the setting where A , A

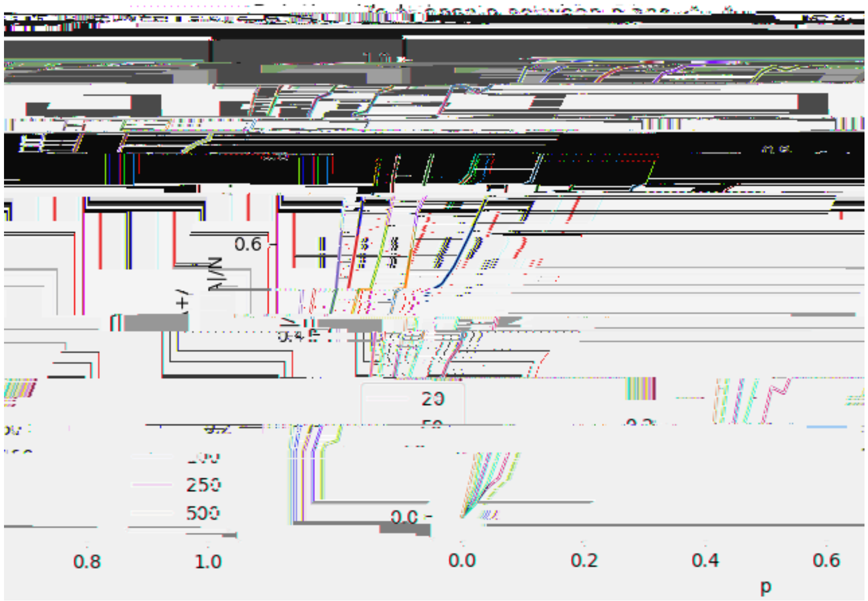

Figure 1: Relationship between p and $|A+A|$

Now we can analyze this problem analytically and check if this is consistent with what we have in our simulation.

Theorem 1. *Let $A \subseteq \mathbb{Z}_N$ be a randomly chosen subset where each element of $\mathbb{Z}_N$ is independently chosen to be in A with probability p. If N is odd, then*

$$E|A+A| = N\left(1-(1-p)(1-p^2)^{\frac{N-1}{2}}\right)$$

If N is even, then

$$E|A+A| = N\left(\frac{2-(1-p)^2(1-p^2)^{\frac{N}{2}-1}-(1-p^2)^{\frac{N}{2}}}{2}\right)$$

Proof.

For all N, we have

$$E[|A+A|] = E\left[\sum_{i=0}^{N-1} 1_{A+A}(i)\right] = \sum_{i=0}^{N-1} E[1_{A+A}(i)] = \sum_{i=0}^{N-1} P(i \in A+A)$$

Note that $i \in A+A$ if and only if there is some $u;v \in A$ where $u+v=i \mod N$

Therefore, if $i \notin A+A$, it has to be the case that for all $(u;v)$ pair where $u+v=i \mod N$, at least one of them is not in A

If $u \neq v$, then the possibility of at least one of them not in A is $1-p^2$.

On the other hand, if $u=v$, then the possibility of at least one of them not in A is $1-p$

So the problem is essentially reduced to finding the number of such pairs

Suppose N is odd

I claim that for all $i \in \mathbb{Z}_N$, there is exactly one element

Figure 2: Comparison of simulation and analytic result

3 $E[|A + A|]$ in Z

Now whatn

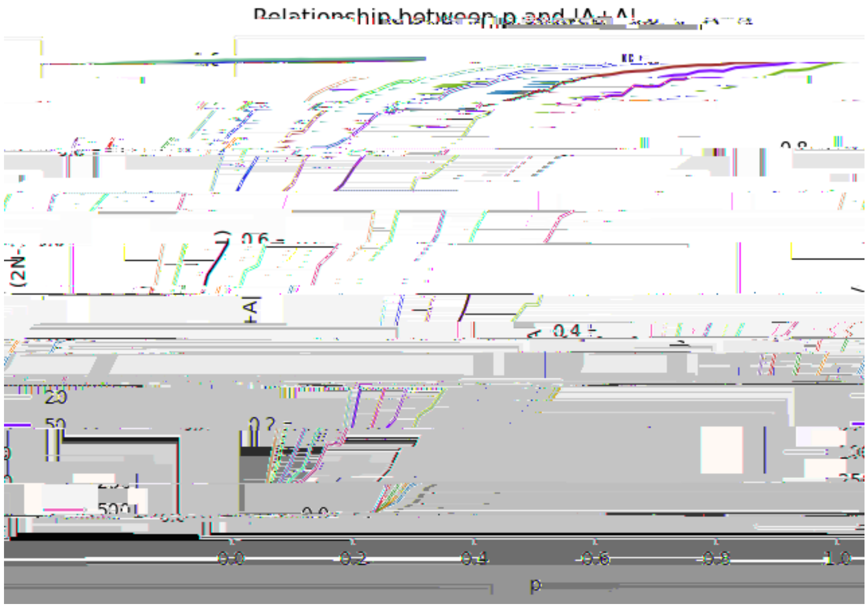

Figure 3: Relationship between p and $|A+A|$

Now we can analyze this problem analytically and check if this is consistent with what we have in our simulation. For simplicity, we will work with the case N being even so exactly half of the numbers in $\{1, 2, \dots, N\}$ are even and half of them are odd

Theorem 2. *Let $A \subseteq \mathbb{Z}$ be a randomly chosen subset where each element of $\{1, 2, \dots, N\}$ is independently chosen to be in A with probability p. If N is even, then*

$$E|A+A| = 2N - (4 - 2p - p^2)\frac{1-(1-p^2)^{\frac{N}{2}}}{p^2}$$

Proof. This time, we have

$$E[|A+A|] = E\left[\sum_{i\in\mathbb{Z}} 1_{A+A}(i)\right] = \sum_{i\in\mathbb{Z}} E\left[1_{A+A}(i)\right] = \sum_{i\in\mathbb{Z}} P(i \in A+A) = \sum_{i=1}^{2N} P(i \in A+A)$$

Note that we are summing from 1 simply for the simplicity of keeping the number of even and odd number the same. $P(1 \in A+A) = 0$ so it won't affect our analysis.

We can still reduce the problem to finding the number of pairs of $(u, v) \in [1, N] \times [1, N]$ s.t. $u + v = i$ and $u \neq v$ and number of $u \in [1, N]$ where $u + u = i$

There are a total of 4 cases for the value of i: whether it is $\leq N$ or $> N$ and whether it is odd or even

Case 1: $i \in [1, N]$ and i is odd

Case 2: $i \in [1, N]$ and i is even

Case 3: $i \in [N+1, 2N]$ and i is odd

Case 4: $i \in [N+1, 2N]$ and i is even

I claim that in each case, the number of (u, v) pairs and u we are looking for are as follows:

Case 1: There are $\frac{i-1}{2}$ distinct $(u;v)$ pair and 0 u
Case 2: There are $\frac{i}{2}-1$ distinct $(u;v)$ pair and 1 u
Case 3: There are $\frac{i-1}{2}-(i-N-1)$ distinct $(u;v)$ pair and 0 u
Case 4: There are $\frac{i}{2}-1-(i-N-1)$ distinct $(u;v)$ pair and 1 u
This claim is denoted as lemma 3 and is proved in the appendix

As before, the possibility of at least one element of the $(u;v)$ pair not in A is $1-p^2$ and the possibility of u not in A is $1-p$
Using this fact and the result of lemma 3, we can plug in the expression like what we did before and get

$$
\begin{aligned}
E[|A+A|] &= \sum_{i=1}^{2N} P(i \in A+A) \\
&= \sum_{i=1;odd}^{N} 1-(1-p^2)^{\frac{i-1}{2}}
\end{aligned}
$$

For part III, we have

$$
\begin{aligned}
III &= \sum_{i=N+1;odd}^{2N} 1-(1-p^2)^{\frac{i-1}{2}-(i-N-1)} \\
&= \sum_{i=1;odd}^{N} 1-(1-p^2)^{\frac{i+N-1}{2}-(i-1)} \\
&= \sum_{i=1;odd}^{N} 1-(1-p^2)^{\frac{N-i+1}{2}} \\
&= \sum_{k=1}^{\frac{N}{2}} 1-(1-p^2)^{\frac{N-(2k-1)+1}{2}} \\
&= \frac{N}{2}-(1-p^2)\sum_{k=1}^{\frac{N}{2}}(1-p^2)^{\frac{N}{2}-k} \\
&= \frac{N}{2}-(1-p^2)\sum_{k=0}^{\frac{N}{2}-1}(1-p^2)^{k} \\
&= \frac{N}{2}-(1-p^2)\frac{1-(1-p^2)^{\frac{N}{2}}}{p^2}
\end{aligned}
$$

For part IV, we have

$$
\begin{aligned}
IV &= \sum_{i=N+1;even}^{2N} 1-(1-p^2)^{\frac{i}{2}-1-(i-N-1)}(1-p) \\
&= \sum_{i=1;even}^{N} 1-(1-p^2)^{\frac{i+N}{2}-1-(i-1)}(1-p) \\
&= \sum_{i=1;even}^{N} 1-(1-p^2)^{\frac{N-i}{2}}(1-p) \\
&= \sum_{k=1}^{\frac{N}{2}} 1-(1-p^2)^{\frac{N-(2k)}{2}}(1-p) \\
&= \frac{N}{2}-(1-p)\sum_{k=1}^{\frac{N}{2}}(1-p^2)^{\frac{N}{2}-k} \\
&= \frac{N}{2}-(1-p)\sum_{k=0}^{\frac{N}{2}-1}(1-p^2)^{k} \\
&= \frac{N}{2}-(1-p)\frac{1-(1-p^2)^{\frac{N}{2}}}{p^2}
\end{aligned}
$$

Putting everything together, we have

$$
\begin{aligned}
E[\square A + A\square] &= I + II + III + IV \\
&= \frac{N}{2}-\frac{1-(1-p^2)^{\frac{N}{2}}}{p^2}+\frac{N}{2}-(1-p)\frac{1-(1-p^2)^{\frac{N}{2}}}{p}+\frac{N}{2}-(1-p^2)\frac{1-(1-p^2)^{\frac{N}{2}}}{p^2}+\frac{N}{2}-(1-p)\frac{1-(1-p^2)^{\frac{N}{2}}}{p^2} \\
&= 2N-(4-2p-p^2)\frac{1-(1-p^2)^{\frac{N}{2}}}{p^2}
\end{aligned}
$$

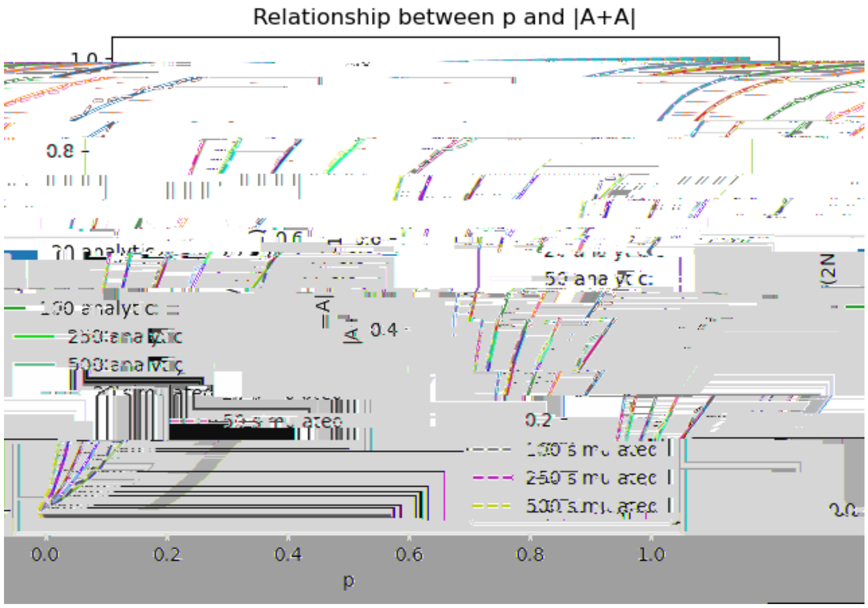

Figure 4: Comparison of simulation and analytic result

We could so similar calculation with the case N being odd and we will have

Theorem 3. *Let $A \subseteq \mathbb{Z}$ be a randomly chosen subset where each element of $\{1, 2, \ldots, N\}$ is independently chosen to be in A with probability p. If N is odd, then*

$$E|A + A| = 2N - (2-p)\frac{1-(1-p^2)^{\frac{N+1}{2}}}{p^2} - (2-p-p^2)\frac{1-(1-p^2)^{\frac{N-1}{2}}}{p^2}$$

The proof is almost identical to the N being even counter part except at handling some index more carefully so it is omitted.

4 $E[A \cdot A]$ in Z_p

Now we want to investigate the expected size of $A + A$ in Z_p and the problem we are trying to solve is as follows

Question 3. *Let $A \subseteq \mathbb{Z}_N$ be a randomly chosen subset where each element of $\mathbb{Z}_N$ is independently chosen to be in A with probability p, then what is the expected size of $|A \cdot A|$?*

Again, let's begin by doing some simulation to have an intuitive understanding. The maximum size of $A \cdot A$ is N so we divide the mean by $2N - 1$ to make the graph on the same scale.

But this time, we need to be careful on the choice of N because Z_N will be a monoid or a group depending on whether N is a prime so when doing experiment, it is better to put in some prime numbers and see if there is any difference

The result is given in Figure 5 for the case $N = 19; 20; 47; 50; 97; 100; 241; 250; 499$ and 500 which corresponds to di erent lines in the graph. We can see that as p goes to 1, the simulated expected size of $|A + A|$ converges to N in all cases. However, we can see that when N is prime, we are having faster convergence rate. For example, the blue line corresponds to the case $n = 19$ and the orange line corresponds to the case $n = 20$. Though $19 < 20$, the convergence rate is faster for the $n = 19$ case. Similar thing is happening for the pair $(47; 50)$, $(97; 100)$, $(241; 250)$ and $(499; 500)$. So instead of having a pure monotonic relationship between N and the convergence rate, we are having two cases, one is N being prime and one is N being composite. In each case, the convergence rate increases as N

Proof.

Note that $p \in (0;1)$ so we have

$$p > p^2$$
$$1 - p < 1 - p^2$$
$$(1-p)^{e(i)}(1-p^2)^{d(i)-e(i)} \leq (1-p^2)^{e(i)}(1-p^2)^{d(i)-e(i)}$$
$$1-(1-p)^{e(i)}(1-p^2)^{d(i)-e(i)} \geq 1-(1-p^2)^{e(i)}(1-p^2)^{d(i)-e(i)}$$
$$1-(1-p)^{e(i)}(1-p^2)^{d(i)-e(i)} \geq 1-(1-p^2)^{d(i)}$$

This is true for all i so we can plug this into the conclusion of the previous theorem and get

$$E[|A \cap A'|] \geq \sum_{i=1}^{N} 1-(1-p^2)^{d(i)}$$

Now we have successfully removed $e(i)$ from our expression and we want to further remove this $d(i)$

I claim that $\sum_{i=0}^{N-1} d(i) = \frac{N(N+1)}{2}$ and this is proven in appendix as lemma 5

To utilize this equation, we can use Taylor Theorem to expand $(1-p^2)^{d(i)}$ as a function of p^2

The underlying function will be $f(x) = (1-x)^{d(i)}$ where $d(i)$ is some fixed constant for now that satisfied the property $d(i) \geq 0$ and $d(i) \in \mathbb{Z}$

Then we have $f'(x) = -d(i)(1-x)^{d(i)-1}$ and $f''(x) = d(i)(d(i)-1)(1-x)^{d(i)-2}$

So we can expand f at 0 and by Taylor Theorem, we have

$$f(x) = 1 - d(i)x + \frac{d(i)(d(i)-1)(1-\xi)^2}{2}x^2$$

for some $\xi \in (0;x)$

So if we plug in $x = p^2$, we have

$$(1-p^2)^{d(i)} = 1 - d(i)p^2 + \frac{d(i)(d(i)-1)(1-\xi)^2}{2}p^4$$

for some $\xi \in (0;p^2)$

Plugging this into the bound we have for $E[|A \cap A'|]$ gives us

$$\begin{aligned} E[|A \cap A'|] &\geq \sum_{i=0}^{N-1} 1 - 1 + d(i)p^2 - \frac{d(i)(d(i)-1)(1-\xi)^2}{2}p^4 \\ &= \sum_{i=0}^{N-1} d(i)p^2 - \frac{d(i)(d(i)-1)(1-\xi)^2}{2}p^4 \\ &\geq \sum_{i=0}^{N-1} d(i)p^2 - \frac{d(i)(d(i)-1)}{2}p^4 \\ &= p^2\frac{N(N+1)}{2} - \frac{p^4}{2}\sum_{i=0}^{N-1} d(i)(d(i)-1) \end{aligned}$$

At this point, the second term involves the sum of $d(i)^2$ which we don't have an estimate on. What we could do is to pull out one of them as $\max d(i)$ and get a bound on this term.

I claim that max

We can see that $d(0)$ is like the outlier that would mess up with our estimation on $\max d(i)$ by a factor of roughly 2. To have a tighter bound, we split the sum into two parts

$$
\begin{aligned}
[\ A\ \ A\ ] &\geq p^2\frac{N(N+1)}{2} - \frac{p^4}{2}d(0)(d(0)-1) - \frac{p^4}{2}\sum_{i=1}^{N-1} d(i)(d(i)-1) \\
&\geq p^2\frac{N(N+1)}{2} - \frac{p^4}{2}N(N-1) - \frac{p^4}{2}\max_{i\ \mathbb{Z}_N; i\neq 0}(d(i))\sum_{i=1}^{N-1} d(i) - 1 \\
&= p^2\frac{N(N+1)}{2} - \frac{p^4}{2}N(N-1) - \frac{p^4}{2}\frac{N+1}{2}\left(\frac{N(N+1)}{2} - N - (N-1)\right) \\
&= p^2\frac{N(N+1)}{2} - \frac{p^4}{2}N(N-1) - \frac{p^4}{8}(N+1)(N^2-3N+2) \\
&= p^2\frac{N(N+1)}{2} - \frac{p^4}{8}(N-1)(N^2+3N-2)
\end{aligned}
$$

□

We could interpret this result as $p^2\frac{N(N+1)}{2}$ is like the main term we are interested in and the $-\frac{p^4}{8}(N-1)(N^2+3N-2)$ is the remainder. We can solve for the range of p where the main term is dominant and the range is $p^2 \leq \frac{4N(N+1)}{(N-1)(N^2+3N-2)}$ which is roughly on the order of $\frac{1}{N}$. This is not a coincidence and is a direct result of our estimation on $\max d(i)$.

Now let's utilize the simulation tool again and see how this lower bound is performing. Since we are only investigating the case when N is a prime, we only check the case when $N = 19; 47; 97; 241; 499$ and compare the simulated result with the lower bound. The result is given in Figure 6 where the right one is a zoomed in graph to see its performance in the low p range more clearly. The lower bounds are all plotted only in the rang $0 \leq p^2 \leq \frac{4N(N+1)}{(N-1)(N^2+3N-2)}$.

We can see that the sharpness of the bound increases as N increases. If we take the same $0 \leq p^2 < \frac{4N(N+1)}{(N-1)(N^2+3N-2)}$ as the range of p, then asymptotically, the remainder will be dominated and the meaning is that every pair of (;) $A\ A$ is given us di erent . But it is also worth noticing that the upper bound of the range of p also goes to 0 asymptotically. So only when we are selecting the elements for A \sparse" and random enough, can we obtain this uniqueness.

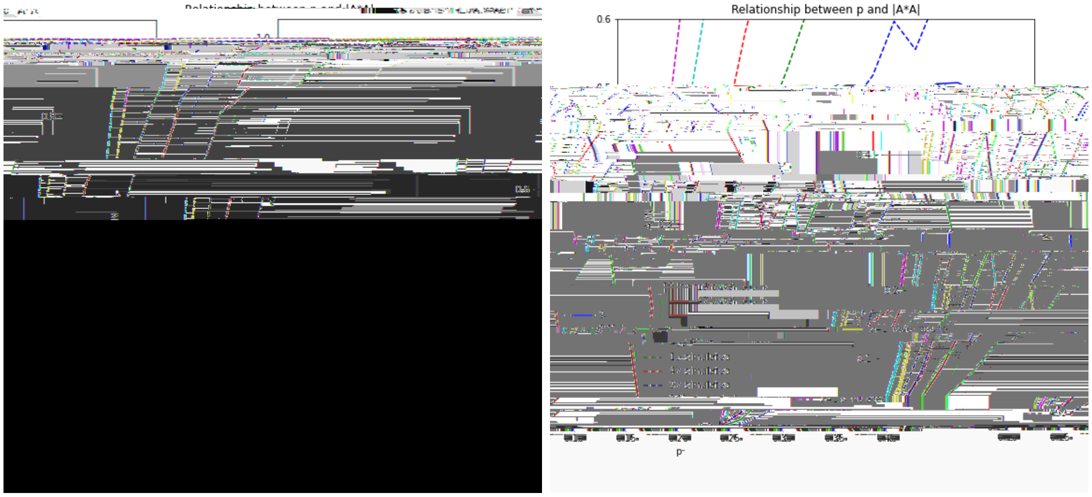

Figure 6: Comparison of simulation and analytical lower bound

5 $E|A \cdot A|$ in $\mathbb{Z}$

The extension of the problem from $\mathbb{Z}_N$ to $\mathbb{Z}$ is very natural and the question we have now is

Question 4. *Let $A \subseteq \{1, 2, \ldots, N\}$ be a randomly chosen subset where each element of $\{1, 2, \ldots, N\}$ is independently chosen to be in A with probability p, then what is the expected size of $|A \cdot A|$?*

Generally, there are many similarities with the above section where we worked in $\mathbb{Z}_N$. Before we dig into the detail, let's first do simulation as always and see what we will get. The range of the product always lies in the range $\{1, 2, \ldots, N^2\}$ so let's just divide the size of simulated $A \cdot A$ by N^2 to normalize the graph as it is unclear from first glance what the maximum size will be. The simulation is done for $N = 20, 50, 100, 250, 500, 1000$ and the result is given in Figure 7.

First, we could see that the ratio is not going to 1. This should be expected as even if all the $(\; ; \;)$ pair gives us distinct product irrespective of order, we should only expect the final size to be $\frac{N(N+1)}{2}$. But even as p approaches 1, this ratio is still lower than 0.5 which suggests that there are some number in the range of $[1, N^2]$ that could never appear in $A \cdot A$. A quick thought will reveal that the prime numbers in the range $(N, N^2]$ will not be in the range. Also, numbers in the same range with a prime divisor that is bigger than N will also not appear in $A \cdot A$. An interesting fact is that the higher N is, the lower the proportion is.

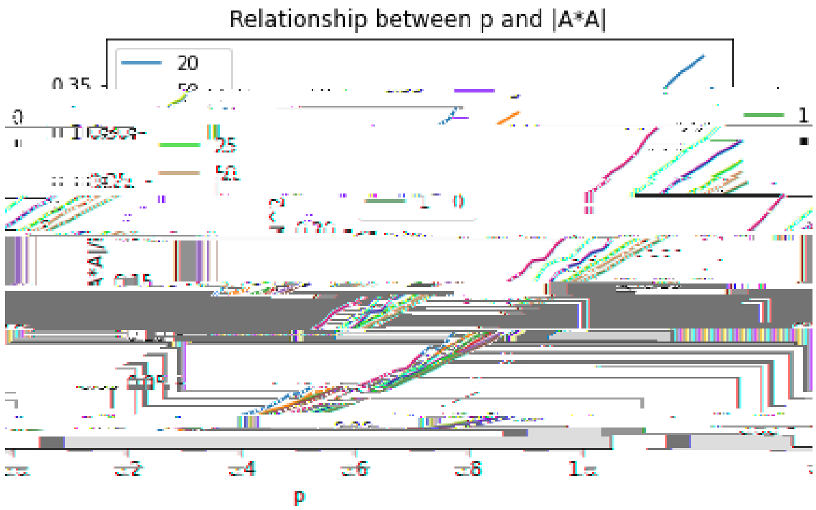

Figure 7: Relationship between p and $|A \cdot A|$

Now we can approach this problem analytically, for simplicity, we again define the corresponding divisor counting function

Definition 5 (Divisor Counting Function in Z)**.** *Let $x \in \{1, \ldots, N^2\}$, define $D(x)$ to be the number of (α, β) pair where $\alpha, \beta \in \{1, \ldots, N\} \times \{1, \ldots, N\}$, $\alpha \leq \beta$ and $\alpha \cdot \beta = x$. This is essentially the number of pairs of divisors of x disregarding the order*

Note that there is no need to define the square root counting function anymore since in Z, a number will either not be a perfect square or it has a unique positive square root, i.e. the square root counting formula is either 0 and 1.

Theorem 6. *Let $A \subseteq \{1, \ldots, N^2\}$ be a randomly chosen subset where each element of $\{1, \ldots, N^2\}$ is independently chosen to be in A with probability p, then*

$$E[|A \cdot A|] = \sum_{i=1; i \text{ not square}}^{N^2} 1 - (1-p^2)^{D(i)} + \sum_{i=1; i \text{ is square}}^{N^2} 1 - (1-p)(1-p^2)^{D(i)-1}$$

Proof.

The proof is essentially the same as the proof of Theorem 4 so I will follow the same logic and be concise

Change the range we are summing, we have

$$E[|A \cdot A|] = E\left[\sum_{i=1}^{N^2} 1_{A \cdot A}(i)\right] = \sum_{i=1}^{N^2}$$

If i is not a perfect square, then its probability of being in $A \cdot A$ is $1-(1-p^2)^{D(i)}$ and if it is a perfect square, then its probability of being in $A \cdot A$ is $1-(1-p)(1-p^2)^{D(i)-1}$

Therefore, splitting into two cases, we have

$$E[|A \cdot A|] = \sum_{i=1; i \text{ not square}}^{N^2} 1-(1-p^2)^{D(i)} + \sum_{i=1; i \text{ is square}}^{N^2} 1-(1-p)(1-p^2)^{D(i)-1}$$

Again, we may face the problem of 0^0 when $p=1$ and as before, we define it to be 1

□

We could approach similarly as in the above section and obtain a cleaner lower bound

Theorem 7. *Let $A \subseteq \{1;\dots;N^2\}$ be a randomly chosen subset where each element of $\{1;\dots;N^2\}$ is independently chosen to be in A with probability $p \in (0;1)$, then*

$$E[|A \cdot A|] \geq p^2 \frac{N(N+1)}{2} - p^4 \frac{o(N^\epsilon)-1}{4} N(N+1)$$

where $o(N^\epsilon)$ means that in little o notation, this quantity is $o(N^\epsilon)$ for all $\epsilon > 0$, i.e. it is "sub-polynomial"

Proof.

Again, the proof will be similar to that of theorem 5.

Since $p \in (0;1)$, we have

$$p > p^2$$

$$1-p < 1-p^2$$

$$(1-p)(1-p^2)^{D(i)-1} \leq (1-p^2)(1-p^2)^{D(i)-1}$$

$$1-(1-p)(1-p^2)^{D(i)-1} \geq 1-(1-p^2)(1-p^2)^{D(i)-1}$$

$$1-(1-p)(1-p^2)^{D(i)-1} \geq 1-(1-p^2)^{D(i)}$$

So we can merge the two sums in Theorem 6 and get

$$E[|A \cdot A|] \geq \sum_{i=1; i \text{ not square}}^{N^2} 1-(1-p^2)^{D(i)} + \sum_{i=1; i \text{ is square}}^{N^2} 1-(1-p^2)^{D(i)} = \sum_{i=1}^{N^2} 1-(1-p^2)^{D(i)}$$

I claim that in this setting, we still have $\sum_{i=1}^{N^2} D(i) = \frac{N(N+1)}{2}$ and this is proved in appendix as lemma 7

Then we again use Taylor theorem on $f(x) = (1-x)^{D(i)}$ and get

$$\begin{aligned} E[|A \cdot A|] &\geq \sum_{i=1}^{N^2} 1-1+D(i)p^2 - \frac{D(i)(D(i)-1)(1-\epsilon)^2}{2} p^4 \\ &= \sum_{i=1}^{N^2} D(i)p^2 - \frac{D(i)(D(i)-1)(1-\epsilon)^2}{2} p^4 \\ &\geq \sum_{i=1}^{N^2} D(i)p^2 - \frac{D(i)(D(i)-1)}{2} p^4 \\ &= p^2 \frac{N(N+1)}{2} - \frac{p^4}{2} \sum_{i=1}^{N^2} D(i)(D(i)-1) \end{aligned}$$

We still want an estimate of $\max_{i \{1,...,N\}} D(i)$ to pull one of the $D(i)$ out.

It is shown in Apostol's textbook [13] that the number of divisor $\ (n)$ is $o(N\ )$ so we should investigate the relationship between the number of divisors $\ (n)$ and our $D(i)$

First note that $D(i)$ is counting pairs of divisors regardless of order so it will be at most be $\frac{\ (n)}{2}$

Moreover, we are restricting our attention to divisor that is smaller than N so this will further reduce $D(i)$ by some unknown quantity so it is safe to use $\frac{\ (n)}{2}$ as our upper bound for $D(i)$

Since little o notation doesn't care constant, we could directly write the upper bound of $D(i)$ as $o(N\ )$

It is worth noticing that when plugging in the bound, we can't pull out the $D(i)$ as we did before since $\sum_{i=1}^{N^2} D(i) - 1$ will then be a negative number

So we pull out the $D(i) - 1$ instead and get

$$
\begin{aligned}
[\ A \ A\ ] &\geq p^2 \frac{N(N+1)}{2} - \frac{p^4}{2} \sum_{i=1}^{N^2} D(i)(D(i)-1) \\
&\geq p^2 \frac{N(N+1)}{2} - \frac{p^4}{2} \max_{i \{1,...,N^2\}} (D(i)-1) \sum_{i=1}^{N^2} D(i) \\
&= p^2 \frac{N(N+1)}{2} - \frac{p^4}{2} (o(N\ ) - 1) \frac{N(N+1)}{2}
\end{aligned}
$$

□

It will be harder to visualize the result due to the fact that the little o notation is not a very precise notation. The bound on $\ (n)$ is more precisely calculated by Wigert[14] to be $\limsup_{n\to\infty} \frac{\log\ (n)}{\log n / \log\log n} = \log 2$. This means that the maximum of $\ (n)$ is on the order of $e^{\frac{\log(2)\log(n)}{\log(\log(n))}}$. As argued above, our upper bound for $D(n)$ is $\frac{\ (n)}{2}$ so I will substitute $\frac{1}{2}e^{\frac{\log(2)\log(n)}{}}$

2 the >N
So 2430548 2430932 T d 65445 2 ang 45 215

log(log(log (n)

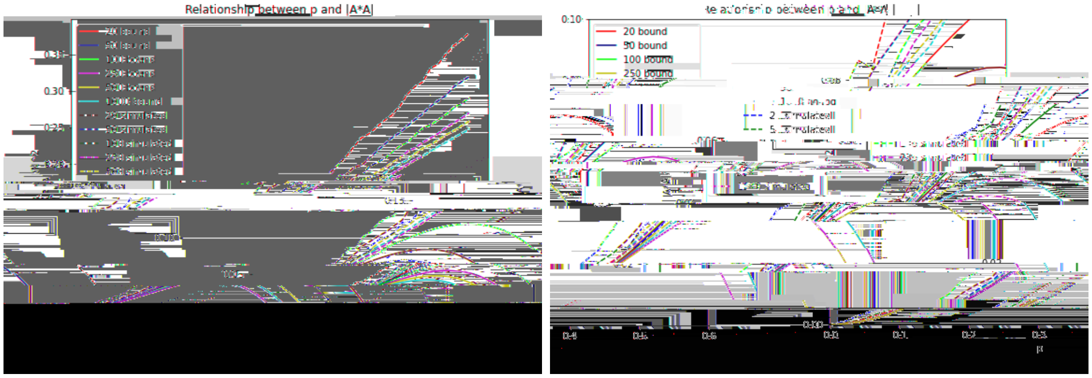

Figure 8: Comparison of simulation and analytical lower bound

6 Appendix

6.1 Proof of Lemmas

Lemma 1. *Suppose* N

However, since i is odd and N is even, the RHS is always odd but the LHS is even so neither could happen

So there are no u s.t. $u + u = i \mod N$

Now I want to show that there are $\frac{N}{2}$ distinct pairs of $(u; v)$ s.t. $u + v = i \mod N$

Note that for a give u, there is a unique v s.t. $u + v = i \mod N$ since Z_N is a group under addition and we just showed that there is no pair where $v = u$

So there will be N pairs of $(u; v)$ that satisfy the condition $u + v = i \mod N$ but $u \neq v$

But our uniqueness definition allows switching the order of $u; v$ so we have to divide this number by 2 and we have that there are $\frac{N}{2}$ distinct pairs

Then, we consider the case i is even

Our previous proof have shown that there are a maximum of 2 elements where $u + u = i \mod N$ since either $2u = i$ or $2u = i + N$

In the current case, both $\frac{i}{2}$ and $\frac{i+N}{2}$ are in Z_N so we are done showing that there are exactly 2 u where $u + u = i \mod N$

Finally, I want to show that there are $\frac{N}{2} - 1$ distinct pairs of $(u; v)$ s.t. $u + v = i \mod N$

Note that for a give u, there is a unique v s.t. $u + v = i \mod N$ since Z_N is a group under addition and we just showed that there are 2 pair where $v = u$

So there will be $N - 2$ pairs of $(u; v)$ that satisfy the condition $u + v = i \mod N$ but $u \neq v$

But our uniqueness definition allows switching the order of $u; v$ so we have to divide this number by 2 and we have that there are $\frac{N}{2} - 1$ distinct pairs

□

Lemma 3. *If N is even, then*

1. *If $i \in [1; N]$ and i is odd, then there are exactly $\frac{i-1}{2}$ distinct $(u; v)$ pairs s.t. $(u; v) \in [1; N] \times [1; N]$, $u + v = i$ and $u \neq v$ and no $u \in [1; N]$ s.t. $u + u = i$*

2. *If $i \in [1; N]$ and i is even, then there are exactly $\frac{i}{2} - 1$ distinct $(u; v)$ pairs s.t. $(u; v) \in [1; N] \times [1; N]$, $u + v = i$ and $u \neq v$ and 1 $u \in [1; N]$ s.t. $u + u = i$*

3. *If $i \in [N + 1; 2N]$ and i is odd, then there are exactly $\frac{i-1}{2} - (i - N - 1)$ distinct $(u; v)$ pairs s.t. $(u; v) \in [1; N] \times [1; N]$, $u + v = i$ and $u \neq v$ and no $u \in [1; N]$ s.t. $u + u = i$*

4. *If $i \in [N + 1; 2N]$ and i is even, then there are exactly $\frac{i}{2} - 1 - (i - N - 1)$ distinct $(u; v)$ pairs s.t. $(u; v) \in [1; N] \times [1; N]$, $u + v = i$ and $u \neq v$ and 1 $u \in [1; N]$ s.t. $u + u = i$*

Proof.

This is a simple book keeping exercise

1. If $i \in [1, N]$ and i is odd, then clearly there is no u where $u + u = i$ since $2u$ has to be even

   Since we are considering Td 0 Td [(])]TJ/F19 3llf 8.u

N being odd prime

This shows that $e(i) \geq 2$

Now suppose y is another solution, i.e. $y^2 = i \mod N$

Then $x^2 - y^2 = 0 \mod N$

Since $\mathbb{Z}_N$ is abelian, we have $(x-y)(x+y) = 0 \mod N$

Since N is odd prime, we know that there is no zero divisors so either $x - y = 0 \mod N$ or $x + y = 0 \mod N$

In the first case, we have $x = y$ and in the second case we have $x = -y$ so y is already counted

So there will be only 2 solutions which is x and $-x$

□

Lemma 5. *Consider the divisor counting Function $d(x)$ in the setting of $\mathbb{Z}_N$, then $\sum_{i=0}^{N-1} d(i) = \frac{N(N+1)}{2}$*

Proof.

Note that $\sum_{i=0}^{N-1} d(i)$ is essentially summing up the number of divisors of all the elements in $\mathbb{Z}_N$

All pair (;) where ; ∈ $\mathbb{Z}_N$ and ≤ , we will have ∣ ∈ $\mathbb{Z}_N$ so it will be counted exactly once when we sum up all the divisor pairs

So $\sum_{i=0}^{N-1} d(i)$ is essentially the number of (;) pairs where ; ∈ $\mathbb{Z}_N$ and ≤

This is a simple combinatoric problem and the answer is $\frac{N(N+1)}{2}$ so we are done □

Lemma 6. *In the setting of $\mathbb{Z}_N$ where N is an odd prime and $d(x)$ is the divisor counting function, then $d(0) = N$ and* $\max_{i \in \mathbb{Z}_N; i \neq 0} d(i)$ where

Lemma 7. *Consider the divisor counting Function $D(x)$ in the setting of $\mathbb{Z}$, then $\sum_{i=1}^{N^2} D(i) = \frac{N(N+1)}{2}$*

Proof.
Similar to lemma 5, $\sum_{i=1}^{N^2} D(i)$ is essentially summing up the number of pairs of $(\ ; \ )$ where $\ ; \ \in \{1; \dots ; N\}$, $\ \leq \ $ and $\ \leq N^2$
This is again a simple combinatoric problem and the answer is $\frac{N(N+1)}{2}$ so we are done □

6.2 Simulation Code

All the simulation done in this thesis is written in Python. In all four cases, the code has high similarity besides from the operator used (plus or times) and the setting (whether need to calculate mod N) so I only included the $A \square A$ in $\mathbb{Z}_N$ as an example. This is a paralleled version that speeds up the simulation by running the program across cpu cores simultaneously.

```
# -*- coding: utf-8 -*-
"""
Created on Sat Apr  3 14:50:59 2021

@author: danny
"""


import numpy as np
from scipy.spatial.distance import pdist
import pandas as pd
from matplotlib import pyplot as plt
from multiprocessing import cpu_count
from multiprocessing import Pool
import os


class Function:
    def __init__(self, N):
        self.N = N


    def get_size(self, p):
        result = []
        for i in range(100):
            chosen_index = np.random.binomial(1, p, self.N)
            A = np.unique(chosen_index * range(1, self.N+1)) - 1
```

```
            A = A[A>=0]
            AplusA = np.unique(pdist(A.reshape((-1,1)),metric = lambda x,y: x*y)) % self.N
            AplusA = np.append(AplusA, A**2 % self.N)
            AplusA = np.unique(AplusA)
            result.append(len(AplusA))
        print(p)
        return np.mean(result)
def main():
    N_list = [19,20,47,50,97,100,241,250,499,500]
    plt.figure()
    for N in N_list:
        p_list = np.linspace(0,1,50)
        if not os.path.isfile("Data/AtimesA_"+str(N)+"_result.csv"):
        resulf.appendsult.append(N)1,50)starma0-0-525(=)-.get_size*2
        fesulfigarpendsjlt:appendvb9d.esA_Fram
```

References

[1] P. Erd}os and E. Szemeredi, \On sums and products of integers," in *Studies in pure mathematics*. Springer, 1983, pp. 213{218.

[2] M. Nathanson, \On sums and products of integers," *Proceedings of the American Mathematical Society*, vol. 125, no. 1, pp. 9{16, 1997.