Volume 45 Spring 2023

MURJ

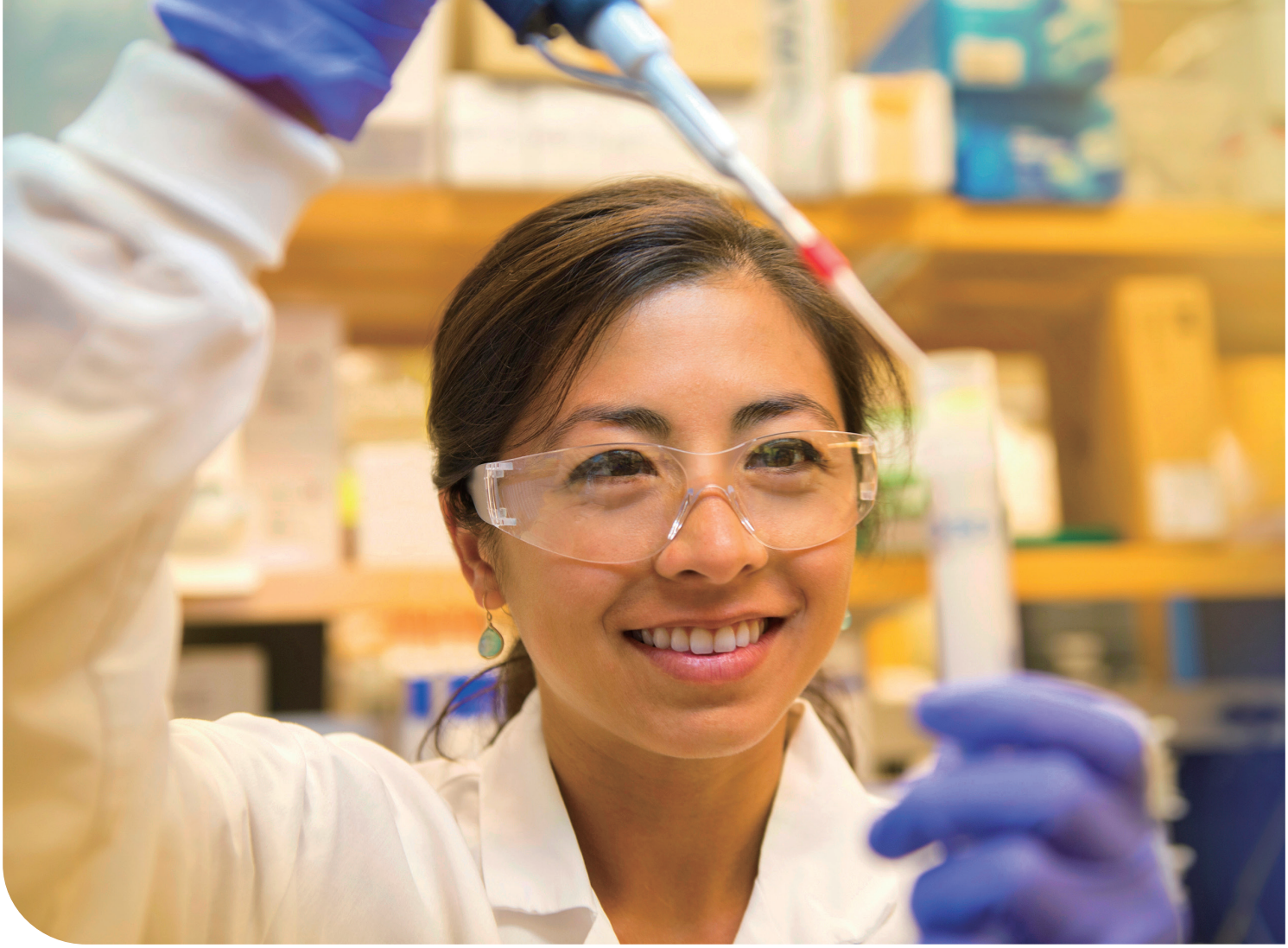
Massachusetts Institute of Technology
Undergraduate Research Journal

News p. 5

Welcome to the Big Players Only Stage

Features p. 12

A Deep-Dive Into Mathematics Research at MIT



Better Health, Brighter Future

Takeda is a global, R&D-driven biopharmaceutical company committed to discovering and delivering life-changing treatments and vaccines that have a lasting impact on society.

Since our founding in 1781 in a market stall in Osaka, Japan, our values endure by putting patient needs first, building trust with society, strengthening our reputation, and developing the business - in that order.

www.takeda.com

Takeda

Contents

Introductory Letter

2 From the Editor

News

5 **Welcome to the Big Players Only Stage**
Arbri Kopliku

Features

10 **Cubic Boron Arsenide: The Semiconductor of the Future**
Rishab Parthasarathy

12 **A Deep-Dive Into Mathematics Research at MIT**
Gwyneth Margaux Tangog and Hannah Han

UROP Summaries

24 **MobilePath: A Mobile Server Platform for Health and Behavior Monitoring**
Sabrina Do, Valerie Kwek, Somaia Saba, Ariadne Dulchinos, Richard Fletcher

27 **Electrowetting Valve Manufacturing and Testing for Electrospray Thruster Production**
Nigel Barnett, Matthew Corrado, Paulo Lozano

Reports

31 **On Symmetric Rank Decompositions of the 3x3 Matrix Multiplication Tensor**
Jason Yang, Virginia Vassilevska Williams

Massachusetts Institute of Technology

UNDERGRADUATE RESEARCH JOURNAL
Volume 45, Spring 2023

MURJ Staff
MIT Undergraduate Research Journal

Massachusetts Institute of Technology

May 2023

Dear MIT Community,

We are incredibly excited to share the 45th issue of the MIT Undergraduate Research Journal (MURJ). MURJ is a biannual, student-led publication featuring innovative undergraduate research at MIT. In this issue, we are proud to showcase the research contributions of students at MIT, which are the product of curiosity, perseverance, and a drive for excellence.

Research in this issue highlights the integration of science and engineering with technology. Advancements in artificial intelligence and computing have greatly accelerated scientific progress and interdisciplinary collaboration. From developing a mobile platform for physicians to monitor health behavior to optimizing electrospray thrusters for satellite propulsion, the work of students reflects their relentless pursuit to further knowledge and engineer solutions at the cutting-edge of engineering and technology.

We are also pleased to share wisdom imparted by professors and alumni in industry. Retiring linear algebra professor Gilbert Strang takes a deep-dive on mathematics research at MIT while Angela Koehler, associate director of the Koch Institute for Integrative Cancer Research, sheds light on the state of affairs of the fast-paced biotech research ecosystem. The breadth of perspectives in biology, mathematics, and computation not only enriches our understanding of these areas, but enables us to draw synergies between adjacent fields.

As research is seldom a solitary task, it would be amiss to acknowledge the contributions of the authors and interviewees in this issue. The publication of this journal is an undertaking by a team of dedicated students and this issue would not have been possible without the hard work of MURJ research, content, and layout staff. We hope this issue of MURJ and the work of its contributors is both enlightening and inspiring.

Warm regards,

Anusha Puri
Editor-in-Chief

abbvie

Would you like to contribute and improve people's lives?

We Offer That.

We leverage more than 130 years of innovation with therapies in 33 disease areas, which means we're creating limitless ways to make an impact. We're extraordinarily passionate about our work but we also know how deeply meaningful it is to cultivate a fulfilling life outside of the lab.

That's what makes AbbVie such a perfect career fit.

People. Passion. Possibilities.®

Explore opportunities and find your fit at **abbvie.com/careers**

MilliporeSigma

Power your progress with Milli-Q® Lab Water Solutions

Your science relies on consistent Lab Water quality. And that needs innovation.

At Milli-Q® Lab Water Solutions, we are inspired by your work to go above and beyond in creating products and services of consistent quality that you can always depend on, now and in the future. We can help you adapt to challenges that lie ahead with our award-winning innovations, and help reduce your environmental impact through sustainable solutions – so you can focus on what matters.

Milli-Q® Lab Water Solutions can give you the confidence of having consistent water quality each and every time.

Power your scientific progress with the Milli-Q® portfolio at SigmaAldrich.com/milliq

40990 04/2022

MilliporeSigma is the U.S. and Canada Life Science business of Merck KGaA, Darmstadt, Germany.

Milli-Q®
Lab Water Solutions

Biotechnology

Welcome to the Big Players Only Stage

A brief commentary on the current state of affairs in the biotech research ecosystem when innovation is at stake, narrated by its protagonists.

By Arbri Kopliku

"*The biotech market is currently oversaturated with companies that simply went public on great ideas, and we are now entering the big players only stage,*" said Dr. Angela Koehler, Associate Professor of Biological Engineering at MIT, in an exclusive interview in late February. She was sharing her perspective on the state of innovation in the biotech environment as the founder of multiple startups in parallel to her academic work, and was thus providing an answer to the question that inspired my commentary piece published in the Fall 2021 edition of MURJ. What makes Professor Koehler's observation so insightful however is that on April 1st 2023, merely two hours before I started drafting this follow-up piece, it was announced that Dr. Samarth Kulkarni, CEO of CRISPR Therapeutics, sold 1.2M$ worth of his company stocks for the third month in a row.

While this news by itself might not spell doom for CRISPR Tx. as a company, when combined with that of the stepping down of Dr. Brad Bolzon, senior advisor in the CRISPR Tx. board of directors, and with a backdrop of stock prices halving in value since last July, potential investments in the company surely do not seem as promising. Is this an unlucky coincidence, or does the case of CRISPR Tx. present an opportunity to cut deeper into the inner workings of the post-COVID-19 biotech market?

One could reasonably argue for the former by pointing out that Dr. Emmanuelle Charpentier, 2020 Nobel Laureate in Chemistry, and Dr. Craig Mello, 2006 Nobel Laureate in Physiology or Medicine, are company co-founders and currently serve on its Scientific Advisory board, while the CEO Dr. Kulkarni himself was a heavyweight of venture capital before joining CRISPR. After all, if this star-studded C-suite and the premise of innovative genetic medicines targeting hemoglobinopathies, immuno-oncology, and tissue regeneration is not enough to grant the title *big player,* what is? According to Dr. Koehler, the real factor governing the ability of companies to raise enough funds to stay afloat in the aggressive biotech market of today is not the brainstorming potential of established PIs willing to lock themselves in discussion, but demonstrated *in vivo* efficacy and the readiness to market products. Indeed, only one of the seven projects in CRISPR's "*In Vivo* Approaches" pipeline has made it to the Investigational New Drug (IND-enabling) stage, and none of them have reached clinical trials. Even CTX001, the first major gene editing project of the company and what Dr. Kulkarni referred to as "practically a cure for Sickle Cell Disease and β-Thalassemia," is only scheduled to finish its Phase 2 trials in late 2024.

A major component of the biotech environment that has changed since the publication of the Fall 2021 piece, and whose significance must be taken into consideration before committing to generalizations from case studies, is that the industry simply has to make do with less capital. The pandemic-fueled flow of investments into biotech from other conventional but less-promising sectors began ceasing in the summer of 2022, and the NASDAQ Biotechnology Index (a rough proxy of how well the biopharmaceutical industry is performing as a whole) has decreased by more than 25% in share price since October 2021. The collapse of the Silicon Valley Bank on March 10th of this year stunned many early-stage biotech startups, while the more weathered ones are constantly having to re-emphasize the relevance of their mission and progress. Dr. Josh Speros, Investment Manager at the Venture Capital arm of BASF, a major biochemical manufacturer, proudly mentioned in an interview that he and his team of four went over more than 1500 funding applications submitted to BASF over the past year: all with incredible ideas, some with finished products, even fewer with patents protecting those inventions. Only 50 of those companies ended up receiving funding from BASF.

Many biopharmaceutical companies that rode the COVID-19 wave of unprecedented cash influx despite not having COVID-19-related products saw their stock values dwindle in the recent biotech crash which visibly shook stakeholders, but a less appreciated consequence is the impact that this drop has had on the senior executive level,

Dr. Samarth Kulkarni (red arrow), CEO of CRISPR Therapeutics, after a talk with the MIT Biotech Group.

says Professor Koehler. Simply put, if you were hired in the immediate post-pandemic biotech market and offered a generous amount of stock options with prices that at the time seemed conservative, chances are that with the devaluation of biotech companies your stock options might be more expensive than the current market value, putting you *underwater*. Without any major stakes in the company, there is not much keeping these C-suite executives from a fresh start with a new company and an exciting prospect, which raises concerns among the HR teams of the *big players* in parallel to promoting innovation and optimizing marketability.

A name certainly affiliated with the biggest of players is Arpa Garay, Chief Commercial Officer at Moderna and MIT alumna. After 16 years of executive positions at Merck, including President of Global Oncology and Chief Marketing Officer, she joined Moderna in May 2022 and brought with her a business-minded approach that promised to turn a rapidly expanding company with a startup mentality into an established biotech actor. Her work has shown that innovation in biotech does not have to come from the lab: we live at a time when the opinion of the customers overrides that of the doctors, and the commercial practices of any company willing to succeed should follow suit. After creating entire commercial teams from scratch, under Mrs. Garay's leadership Moderna secured partnerships with governments worldwide to build manufacturing facilities in different countries in exchange for fixed epidemic-ready batch vaccine orders with priority delivery. This way, the governments have their health security guaranteed due to the host-priority of the product of each facility, while Moderna locks in predictable contracts to fine-tune their manufacturing capacities. An interesting outcome of this entire project is that to this day Moderna does not have a single salesperson in the field, let alone dedicated campaigns to convince doctors of the worth of new treatments.

Mrs. Garay rightfully admitted that the capacity to innovate in high level marketing approaches is a luxury that not many companies have. Lexington Medical, a small medical device company founded in 2013 and headquartered in Bedford, MA, that specializes in surgical stapler manufacturing, almost had to permanently close its doors due to the COVID-19 pandemic lockdown. Its reliance on conventional face-to-face interactions for arranging shipment orders meant that a disruption in global travel and supply chains could cripple its business model and put the future of the company in danger, despite the quality of its sole product being notably better than the competition in effectiveness, safety and cost. The solution that the leadership team adopted when this worst-case scenario became reality was an aggressive hiring campaign that led to salespersons outnumbering engineers by more than an order of magnitude, and that entire staff is still part of the company today. While Lexington Medical successfully weathered the

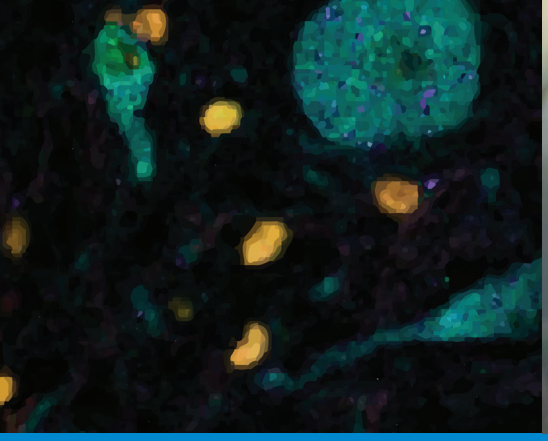
challenges due to COIVD-19 by a mix of fortune and daring hiring decisions, one could see how the company might not be as lucky when the next global market disruption strikes.

We have thus built a somewhat-grim panorama of the post-pandemic biotech industry, and it is ripe time for a discussion of the role that this panorama plays in innovation by looking at two key observations. For the first one, in what sense do Moderna and Lexington Medical differ from CRISPR Tx, that could have influenced their differing prospects for the near future? The answer is intellectual property (IP), and it should come as no surprise given Moderna's openly relentless crusade to hermetically safeguard every platform process component from mRNA medicine composition to delivery mechanisms, even if that means legally challenging Pfizer, Alnylam, BioNTech and the US Government. The case of Lexington Medical was not as eventful, but they did make a point of focusing on a single, well-characterized piece of intellectual property that is their surgical stapler, and of taking advantage of the regulatory leniency by adding custom-made modifications to order that do not need to be re-screened for safety and efficacy by the FDA. When asked about his opinion on the relevance of IP, Dr. Kulkarni, the CRISPR Tx. CEO, answered: "I see people founding successful companies without any IP all the time. While in Europe IPs can be necessary to enter the drug market, here in the US they only serve for peace of mind. After all, licensing technology from a major institution like UC Berkeley or MIT is not that expensive, and provides a reliable business model". That question was a charged one, as MIT and the Broad Institute had won their legal case against UC Berkeley for ownership of the CRISPR invention patent merely 6 months before Dr. Kulkarni's comments. Companies like Intellia and CRISPR Tx, which had been licensing the technology as the foundation of their entire companies from UC Berkeley had just seen their licenses revoked.

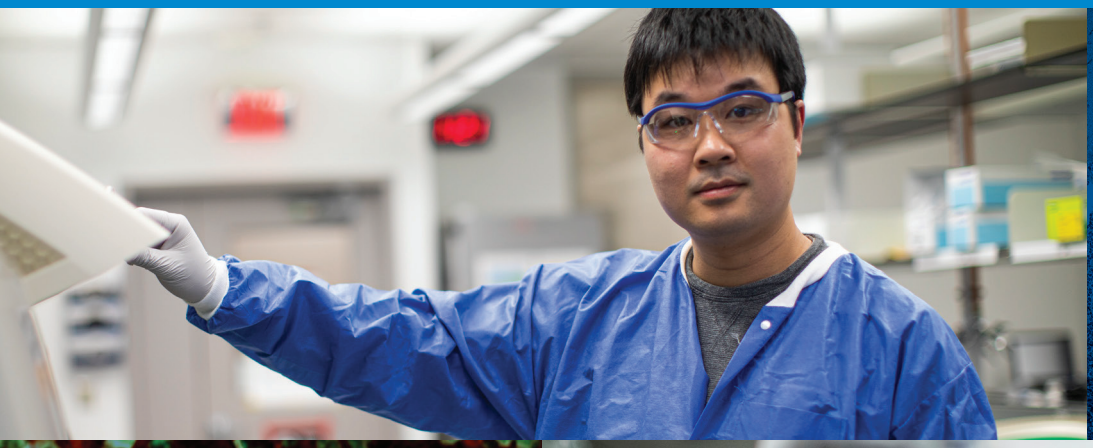
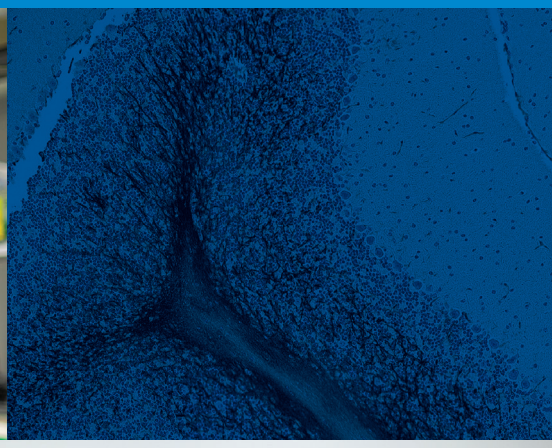
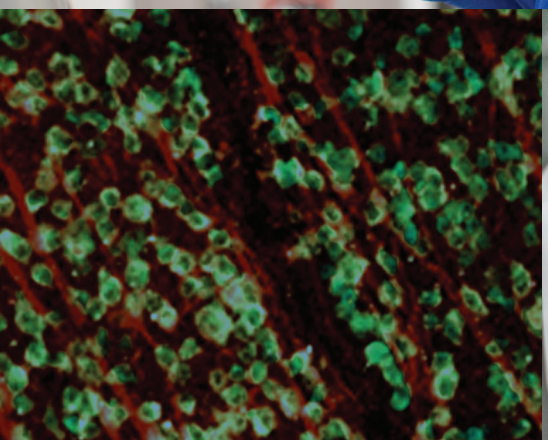
The second observation is that, in stark contrast to the Fall 2021 commentary, from an academic perspective the biotech industry is very much not *anonymous*. A great example of this new visibility is NORA, the Northeast Research Alliance, which is a strategic research initiative led by BASF in collaboration with MIT, Harvard University, and UMass Amherst. This collaboration led to 56 sponsored projects, resulting in 41 patent applications and 27 publications. While the program has existed since 2012, its reach into academia was extended when BASF executives engaged in talks with undergraduates from each of these universities to evaluate the possibility of direct sponsorship for undergraduate biotech-adjacent projects.

"Instead, my interviews with established professor-entrepreneurs, venture capitalists, company CEOs and government agency representatives served as proof that the fundamentals of innovative progress, including patents and the conditionality of funding, apply despite the hostility of the current market towards new players."

This discussion of the challenges imposed by a demanding biotech market upon commercial and academic actors was not meant to confirm or deny whether we really are living in the *roaring 20s of biological research* as proclaimed by Dr. Kulkarni in his talk with the MIT Biotech Group. Instead, my interviews with established professor-entrepreneurs, venture capitalists, company CEOs and government agency representatives served as proof that the fundamentals of innovative progress, including patents and the conditionality of funding, apply despite the hostility of the current market towards new players. The resulting bottle-neck into the *big players only* stage comes with its pros, as can be seen by Moderna's innovation outside the lab to solidify its public image association with an exclusive product of high quality, as well as cons, as demonstrated by the hundreds of potentially game-changing biotech startups whose future is left on the hands of a handful of venture capital investment managers at a giant biochemical conglomerate. Once again, this piece would not have been possible without the candid remarks of all my interviewees and their patience as we carefully traversed a complicated landscape of disruptors and innovators.




where science meets humanity™


Innovation inspired by the passion of our people

As pioneers in neuroscience, Biogen discovers, develops and delivers worldwide innovative therapies for people living with serious neurological diseases as well as related therapeutic adjacencies. We are focused on advancing the industry's most diversified pipeline in neuroscience that will transform the standard of care for patients in several areas of high unmet need.

biogen.com

Biogen®

MURJ Features

Cubic Boron Arsenide: **The Semiconductor of the Future**

Featuring MIT Professor Gang Chen

By Rishab Parthasarathy

Inside all our phones, our computers, even our refrigerators, lies a hidden hero, the silicon microchips that drive all technology worldwide. All this is possible thanks to two properties of silicon: its abundance in the form of silicon dioxide, or sand, and its properties as a semiconductor, a material that conducts electricity worse than conductors (like metals) but better than insulators (like rubber).

However, despite these promising properties, which have led to its ubiquitous use, silicon is anything but perfect. According to MIT Professor Gang Chen, silicon has two large weaknesses. The first is its poor heat conductivity, which is why electronic devices overheat easily. The second is that while silicon conducts electrons well, it lacks the ability to conduct "holes," the negatively-charged electron's positively-charged counterpart.

That's where the idea of cubic boron arsenide arrived, which according to a group of researchers including Prof. Chen and Prof. Zhifang Ren at the University of Houston, solves both of these issues that silicon faces.

According to Prof. Chen, the work began when a frequent collaborator of his, Professor David Broido at Boston College, discovered that cubic boron arsenide was an intriguing material for solving heat issues: "Our initial experimental measurement was heat, … and we worked for a few years and proved that this material really does have high thermal conductivity, and that was 2018." From there, with the experimental properties of cubic boron arsenide verified for the first time, the research took off.

Using research that his group had been separately developing for measuring the electrical properties of materials, Prof. Chen then decided to test cubic boron arsenide through a suite of simulations, discovering that "cubic boron arsenide has both high electron mobility and hole mobility, [which] is very unusual because there are [effectively] no materials" with the same properties.

Now, the new work tries to prove the results in the real world, taking the simulations from numerical estimates to actual measurements. To reach these goals, Prof. Chen began collaborating with Prof. Ren at Houston, who invented the procedure for growing cubic boron arsenide in the lab, a first step towards experimental measurement.

This is where the biggest challenges started. As Prof. Chen stated, "The measurement was hard, because even though we could grow a single crystal, different places had different impurities, … so the material was nonuniform." In essence, the group of researchers could not calculate consistent measurements from the samples of cubic boron arsenide they were able to grow in the lab, leading to significant experimental doubts.

But after coming back to their apparatus for measuring the heat conductivity of cubic boron arsenide, Prof. Chen's team reached a breakthrough. Using the tool originally designed with Prof. Keith Nelson of the MIT Department of Chemistry, the researchers modified the device to send electrical signals through the compound instead of watching heat propagate, using a technique called the "transient gradient

technique" to discover which areas of cubic boron arsenide would yield valuable experimental results. Using this technique, they were finally able to finetune the results which they have now published, verifying that cubic boron arsenide has unique electrical and thermal properties that aren't possessed by any other known semiconductors.

Looking to the future of the technology, Prof. Chen said "[cubic boron arsenide] is more like a cubic material: it's like the same type of material that people currently use for semiconductors," giving cubic boron arsenide a unique advantage over other current avenues of research like graphene. But getting to a place where these materials can replace silicon is still unclear, as "it's still early-stage research showing that these materials have amazing properties".

As for next steps, Prof. Chen hopes that his work has opened the eyes of his fellow researchers to the potential of materials like cubic boron arsenide, a whole class of semiconductors that has not yet been well studied: "It's so simple as a compound, … and it's most exciting that we're demonstrating this potential."

"If you have an [interest or a question], reach out to that professor or the students in that group, and then if it's not at MIT, go beyond. Nowadays, it's so easy to find information, so don't be shy!"

So one day, once industry uses their tools to develop cubic boron arsenide, improving the bottleneck of our ability to efficiently synthesize pure cubic boron arsenide, you might just use these simple compounds in your smartphones of the future, powering your life behind the scenes.

Speaking to undergraduates who want to get into research, Prof. Chen closed by noting the interdisciplinary nature of his work, linking with professors across MIT departments and across national universities. In his words, "if you have an [interest or a question], reach out to that professor or the students in that group, and then if it's not at MIT, go beyond. Nowadays, it's so easy to find information, so don't be shy!"

A Deep-Dive Into Mathematics Research at MIT

Featuring MIT Professors Gilbert Strang, Larry Guth, and Jeremy Hahn

By Gwyneth Margaux Tangog and Hannah Han

The Massachusetts Institute of Technology is known for its rigorous mathematics programs and prolific research department. MIT's accomplished mathematics professors have published research papers in topics ranging from parallel computing to algebraic geometry. However, it can be difficult for students to keep in touch with and learn about the math department's research processes and UROP opportunities, as only a few postings of these are made known to students through platforms such as ELX. To promote mathematics research, MURJ is interviewing a few professors to learn about their experiences in mathematics research, what they are currently working on, and their advice for aspiring mathematicians.

The following are several interview questions answered by Professor Gilbert Strang, Professor Larry Guth, and Professor Jeremy Hahn. Professor Gilbert Strang is MIT's distinguished pioneer in linear algebra and currently teaches 18.06 during his last semester with MIT. His research focuses on mathematical analysis, linear algebra and PDEs (Partial Differential Equations). Professor Larry Guth specializes in metric geometry, combinatorial geometry, and harmonic analysis. Professor Jeremy Hahn joined the MIT department most recently (in 2021); his work specializes in algebraic topology and homotopy theory. The three professors were interviewed about their past and current research, post-undergraduate lives, advice they have for students, and their input on mathematics and technology.

PAST RESEARCH/SIGNIFICANT WORK

Questions asked: "What research/work have you done? Which are your favorite and most impactful?"

Professor Strang

For me, textbooks and video lectures have been a big part; that's probably the most important thing I've done. When I started teaching linear algebra at MIT, 50 or 60 years ago, linear algebra was only for math majors. When I was an undergraduate and I took the course, it didn't have that number 18.06. It was very abstract, very pure, and only a small number of math majors were in it. And to me, as years went on, I thought, 'This is not right!'. Linear algebra is just as important as calculus, and many more students have to know about linear algebra. And it's beautiful! Nice ideas, nice structure. I like it better than calculus.

Calculus has this key idea of slope at a point and area. But linear algebra is coming everywhere now in deep learning, data science, statistics, and many many more applications. So, it was a subject I enjoyed. My work is mostly in textbooks and original articles on linear algebra with applications.

The research I've done is on topics in pure linear algebra: results about the rank, the null space and all the good things, and their applications to engineering and to science generally. I introduced 18.06 for a more basic linear algebra course with more applications instead of all proofs, and it caught on quickly to all departments at MIT. It

Professor Gilbert Strang

Professor Jeremy Hahn

Professor Larry Guth

became a bigger and bigger course, and more fun to teach and more fun to write about. That's my main direction in a nutshell.

My work is always seeing ideas that are either specifically in linear algebra: like facts about the ranks of submatrices and applications in engineering and science where linear algebra comes in. One place that was sort of exciting—I hadn't taken any engineering courses at MIT when I was a student, so I didn't really have a good picture of what some of the problems were. But civil engineers and mechanical engineers needed to solve differential equations. They had devised something called a "finite element method". And it was successful! Engineers were using it, coding it big time, companies were coding the finite element method, and it had many variations. So, naturally, a math person like me says, "What's going on? What's the math behind it?" So that was my part, and a bunch of others also wrote about the finite element method. It's a way to solve differential equations. The original way was to take every derivative and turn it into a finite difference, sort of undo calculus. The finite element method was another way to approach the problem: you invent some trial functions, and the problem would be to find the best combination of the trial functions to solve your equation. It started with trial functions instead of starting with [(y(x+h) - y(x)]. So it was an adventure, and then I wrote a book about what I understood, which was only part of it, and that got me into writing books!

Professor Guth

When I was a postdoc, I got very interested in a problem called the Kakeya problem, which is a geometry problem but it turns out to be related to Fourier analysis and waves and a few different things. The geometry problem is—the original version is—you have a needle that has length one sitting on the table and you want to rotate it all the way around and see how little area it takes to do that. At first it seems like the best way would be to put it inside of a disk and rotate the needle around the disk, but it turns

out that's not the most efficient thing to do. And surprisingly, you can rotate the needle all the way around in an arbitrarily small area – Besicovitch found that in 1920. If instead of being infinitely thin and the needle has some small thickness delta, then it takes a finite amount of area, and then (1/log(δ)) is the most efficient thing to do. That's in two dimensions, which is kind of understood, and then there's a similar problem in higher dimensions. So you have a cylindrical needle radius delta, length one in three dimensions, and you want to have a three dimensional chamber that can hold this needle pointing in any direction. How big does it need to be? Even in spite of Besicovitch's example it seems intuitive and it seems plausible that the area you need is pretty close to the volume of a ball, but nobody knows that, so this problem was around sort of a geometry problem a long time ago but then people realized it's potentially related to questions about wave propagation and questions in Fourier analysis [in terms of] how well Fourier's series model their functions.

Nets Katz[1] and I were very interested in this and we were thinking about different problems that are related to it, and there is a computer scientist who solved a cousin of the Kakeya problem, and he gave this very short proof, two page proof, and it used high degree polynomials in a clever way that was surprising to everyone. And it was based on ideas in error-correcting codes in computer science; they used high degree polynomials. So Nets and I worked together to try to understand how much this could say about the original Kakeya problem. So far, it doesn't seem to solve the original Kakeya problem but we used it to solve some other problems in combinatorics. The problems in combinatorics are about the intersection patterns that you can make, using a finite set of lines or a finite set of circles, whatever sort of basic geometric object. Like if you have a million lines in space, and you'd like to make them intersect a lot, so you make a lot of special start points that have 10 different lines going through them; how should you arrange the million lines to make that happen the most? And there are interesting patterns of lines that come from polynomials, that'd be like some polynomial equation that makes a surface that has a lot of lines in it, if you take the lines in the surface they intersect each other a lot in a nice way. What we took out of Zeev [Dvir]'s[2] proof was a way to detect if that was going to happen, that whenever you have a set of objects that intersect each other a lot, we could find a polynomial that contains them and that sort of makes an algebraic family that explains why they intersect each other so much. So there are fairly many cases where some objects that do something surprising, there's like a polynomial or something from algebra underneath that explains how they do that and there's a method to find that. And then there are lots and lots more questions, where it seems like something like that might be true but we definitely can't prove it.

Professor Hahn

I work in algebraic topology and homotopy theory which I think is a subject that has connections to several different areas. Historically, the idea is to study shapes in high dimensions and try to understand what kinds of shapes are possible and classify shapes up to continuous deformations. But algebraic topology in particular tries to study those through algebraic invariants. Increasingly, in modern years, there's a lot of connections to algebra and a lot of interesting algebra coming out of the subject. What I've done in recent years includes explicitly classifying high-dimensional shapes, a joint work with Robert Burklund and Andrew Senger; classifying high-dimensional shapes satisfying a certain condition is an old problem from the 60s. Also, there's a really exciting development in modern number theory which is exciting, prismatic cohomology. I'm doing a lot of work with a postdoc here to

1 Nets Katz is currently the IBM Professor of Mathematics at the California Institute of Technology. His research interests lie in combinatorics, harmonic analysis, and fluid mechanics.

2 Zeev Dvir is a professor in both the Computer Science and Mathematics departments at Princeton University. His interests include theoretical computer science and mathematics, with special attention to computational complexity, pseudo-randomness, coding theory and discrete geometry.

connect prismatic cohomology with homotopy theory and the structures coming out of algebraic topology.

Probably my most impactful work is something I did in graduate school with fellow student Danny Shi. We managed to connect purely algebraic invariants—known as higher real k theory—to some geometry, geometry of complex bordism. As a result of making that connection, we were able to do a bunch of new computations. That was a lot of fun and seems to have impacted a lot. I think my favorite research changes. Currently, I'm interested in this prismatic cohomology stuff, but I'm sure if you ask that question to me in two years, I would say whatever I'm currently working on.

POST-UNDERGRADUATE LIFE

Questions asked: "What is life like for a mathematician post-undergraduate? Did you ever consider working in industry?"

Professor Strang

Actually, in grad school, all the way through, I had the idea that I would go into industry, without knowing anything that's special about it, except having the idea that you made more money in industry. I had thought, 'Okay, I haven't got any money, I better work somewhere and get some, or my life is gonna be limited.' I was in grad school, I came to MIT as an undergraduate, I went to Oxford for a couple years in England, and then I went to UCLA to finish the PhD. I was at UCLA finishing my PhD, and along came an offer from MIT to be an instructor. MIT has instructors as well as assistant and associate professors. Instructors are usually here for two or three years. It was nice and I liked California, but I really like Boston. So I took that offer. In the end, I changed my idea of being an applied mathematician in industry to being a professor; it turns out that I enjoy teaching! That was a lot of fun; I'm sure that the other options of going to work for IBM or Bell Labs—at that time it was just IBM or Bell Labs, now there are many many companies that hire math people—I'm sure that would've been okay, but I loved this job. Happy I did it. It's been a long time and this is my final semester.

Professor Guth

Well people can do lots of different things. I went to graduate school after graduating college. I did spend a bit of time wondering if I would like to do something else, maybe for a little while. But then I went to graduate school, and that was a big learning time. There were some frustrations but [it was] mostly a very happy time in my life. It was, for me, a lot about learning to read and write. Before that, I hadn't ever read a math paper. I had read some books that were somewhat hard-to-read but, in graduate school, I started spending lots of time reading things, reflecting on them, trying to understand how to read difficult passages and navigate many different books which refer to other books. I spent a lot of time reading things and exploring this new world, and a lot of time writing things. One of the things about UROPs is writing. Almost everyone has this experience the first time that they've thought about something for a while and they've figured out something that makes sense to them. Then, they go to write it down and then they produce something that other people cannot read. [Then we have to ask,] "How do we say things that other people can understand?" My time in graduate school was a lot about practicing those things. It was sweet doing those things with other people. I had a little group of graduate student friends—same adviser—and we'd take turns reading new things and trying to explain them to each other.

"I spent a lot of time reading things and exploring this new world, and a lot of time writing things."

I think that math is useful all over science and engineering, but I also have been thinking that maybe we could do a little better job of preparing people to bridge. I think there are a lot of fields

where it'd be very helpful to have someone who knows some math, but maybe, to work in that field, it's also necessary to know some other things. Sort of figure out how to get people ready for that. Roughly what mathematicians are good at is reading something quite complicated and digesting how it fits together and being able to explain the main points to people and that comes up in all technical things. I never have [worked in industry]. As I've gotten older, it has started to sound more interesting and appealing to me. When I was young, I just enjoyed doing pure math and, as I've gotten older, the things that are a little more applied are maybe making my ears perk up and, someday, [it] sounds really cool to work on something more applied, but I also don't have that much time anymore. When I read the seminar listings or the topics classes in the math department, the ones on math biology or the ones on statistics and data science, they all sound really interesting.

Professor Hahn

From my own personal experience, I went to graduate school. That's different from my life now, but my life in graduate school was one of the happiest periods of my life. I really got to focus on math all day. It's a bit like being an upper-level undergraduate but there's less things to worry about in terms of HASS requirements and other courses and you can focus on things you like. That was a lot of fun. It doesn't pay very well to be a grad student, but it's an extremely nice lifestyle to be able to focus on the subject.

I do pure math here, so unfortunately I haven't had a lot of opportunities to interact with fields outside of mathematics. I think there's a spectrum for pure to applied mathematics and what I see my personal work connecting a lot with is other areas of pure math which might then go on to connect with applied math. I do enjoy seeing things applied in a lot of places. But, personally in my own work, the applications have been to other areas of mathematics. I have a lot of friends, even going back to my undergrad days who have done work in computer science and material science who have previously taken a lot of math with me, and I certainly feel like the math that we took together has influenced some of their work. In particular, computer science is very popular these days, and for good reason. There are connections to a number of areas of math there.

RESEARCH AND GENERAL ADVICE FOR MATHEMATICIANS

Questions asked: "What advice do you have for aspiring mathematicians at MIT? What advice do you have for students looking to do a UROP in the math department?"

Professor Strang

My advice is to look for, sometimes a bunch of problems or an area of math just clicks. You feel you kind of understand it and you see some new questions. That's what you hope for. Hopefully, it's an area which is growing which other people are also interested in. I'm sure there will be more growth. As of now, deep learning is a hot subject, and related problems are, of course, growing quickly. And there are areas of pure math (and applied) which are growing quickly. Have an open mind. But in the end, you have to choose something that clicks with you and which you get [new] ideas about.

Professor Guth

I have done some [UROPS] myself, but not a huge number. I feel like when people I don't know email me, there's more who email me than I can take, I haven't found a solution to that that I feel happy with. I get the impression from a lot of students that I talk with that students feel like a UROP is something totally different from a math class. My personal experience of it is that they're much closer, that kind of thinking when somebody's teaching a class or taking a class, that kind of thinking when they're doing a UROP, are almost the same. The basic situation is that there are some things that we'd like to understand that we don't understand yet, and in class there are other people who have understood those things

at some point, so there are some more clues, but usually in a UROP there are also clues. There are math questions out there where we have no clue what to do, but usually if we give those to a UROP student they also have no clue what to do, and the UROPs that go well are where there are clues. So it's not so different from class, and when we're in class we could allow ourselves to be a little more open about our thoughts; to do the homework but also say, "Well, what do I think is an interesting question about this thing?" and I could try to think about it. Then, I could go to office hours, and say, "Well, I came up with this question and I tried to think about it a little bit but now I'm stuck over here, have you heard anything about that?" I feel that there could be something more [like a] continuum. So, the advice I usually give to undergrads about finding UROPs is to just start thinking that way in their classes and talk to people about things, and it's going to be interesting anyways and probably going to be a good way to find an efficient UROP.

" Have an open mind. But in the end, you have to choose something that clicks with you and which you get [new] ideas about."

It's probably a little trickier to make a good UROP for someone with lower level/introductory classes, but I don't think it's impossible. But the other thing is that, I think in practice, what's happening in the math department is that we have a lot of majors now who want to do math UROPs and it seems like there's only so much professor time and postdoc time. So, given that situation, if there are two people who would like to do a UROP, and one of them has been around longer, it's a little hard to justify not giving the UROP to that person who's taken six classes and worked hard in order to make time to give a UROP [to someone] who's only taken two classes. My impression of the biggest issue is just the total scale, if we're going to be allocating UROPs.

[In terms of other department UROPs], there's lots of interesting stuff going on at MIT. I noticed in the physics department there's someone who was trying to make a large-scale UROP with 15 people as a class; somehow organizing in such a way that everyone was doing research, but everyone had a part of it that was their own. If we want to reach the number of students that we want to, we might try to do things like that. Either [interdisciplinary or just math], but just somehow bigger. They're all interesting; if you try doing a UROP in physics or chemistry or material science or computers then you'll see that there will be some math that comes up in it but it won't be just math, you'll have to also be interested and involved in whatever the subject is. And I think that is pretty similar to what it's like taking some experiments in math and bringing it out into industry or the public sector somewhere and trying to find a job and do something interesting with them.

For doing pure math, I'd say there are a lot of classes that are research-like or have a range of open-ended projects. There's Math Project Lab (18.821) which is about doing research projects and there are a whole range of math seminars that are about doing some kind of reading projects, could be a research project. And almost all of these classes are designed so that they're not a crazy amount of work for everyone who wants to take the class, but they're designed so that you could put into it whatever you want to put into it. For example, in one of the CI-M classes, most of the student's final projects aren't research, but if you wanted to you could. In any case, they are like studying and presenting some relatively recent, interesting development in the subject. Which is very close to what research is. If you want to put a lot of effort into it and digest something that's bigger and more difficult, then you can do that in that class. And the teacher will guide you. That's probably a little bit more like what pure math research is than a UROP in another department. On the other hand if you're interested in doing research where you apply math to other things, then a UROP in another department could be great.

[In general], one thing is that it's great to

explore what you're interested in and to see what you're excited about and what you find that you just want to keep spending time on doing. That's a great thing to see about yourself, that you can go out and appreciate doing X. There are many different signs/sides to the process of that. It could be that you enjoy trying to prove new theorems. It could be that you enjoy learning about what's out there and reflecting about what would be a good question [to ask]. The other thing about research is that there really is a lot of getting stuck and, in classes, the professors would like to challenge the students to some extent but take some responsibility so that they don't give them something where they'll just be stuck for 5 weeks. In research, the typical thing is that we're stuck for long periods of time; it's good to know that everybody goes through it. To get somewhere meaningful in probably any kind of endeavor, we have to go through a lot of being stuck. That's one thing: the emotional side of being stuck. The advice I give most often to my graduate students is to go back and forth between trying to make progress and being stuck, and stepping back and trying to understand, "Why are we stuck? What is the wall we're running into? Describe that wall." I try to do that too. For example, going back to the Kakeya problem, going back and forth between trying to make progress on the Kakeya problem and stepping back and asking, "Why is it that we have all had so much trouble making progress on the Kakeya problem?"

Professor Hahn

I do think that UROPs are very important. When I was an undergraduate here, I was very much influenced by the ability to have UROPs with professors. I had a UROP with Professors Mark Behrens and Clark Barwick over two consecutive years that really interested me and influenced the direction I went my whole life after that. It's really a great thing to have the UROP program here. [However], I understand that getting a UROP is difficult. Every professor has different criteria for what kind of UROP students they take. In algebraic topology, there's a little bit of background necessary to do a worthwhile UROP. What I see students doing is doing directed reading programs with graduate students up until the point where they have a really good background, in particular when they've taken some upper-level graduate classes, which a lot of undergraduates here do towards the end of their time [as an undergraduate]. At that point, I think it's really important for me to offer a good UROP. For my area of math, you need some background. You need to demonstrate that you have that background, [and] I'm always looking for students that have that advanced background in algebraic topology.

Mark Behrens[3] and Clark Barwick[4] who were here when I was an undergraduate—and are no longer here—, they both did homotopy theory, so I still work with them to this day and we are in the same field. It's all because of their generosity with things like UROPs and the classes I took with them that led me to learn the stuff necessary to start those UROPs.

[In general], I understand there's a lot of pressure here in the sense that there are a lot of really strong students. I certainly felt that pressure a lot as an undergraduate here. I think it's important to remember that you can do quite well as a research mathematician. You don't have to necessarily be the fastest or the most prepared. There's a lot of competition or pressure from fellow students here that it's [all the more] important to have some confidence in yourself that if you put in some work you can do well.

MATHEMATICS AND TECHNOLOGY

Questions: "What is the effect of technology on what the current generation of mathematicians have to study now? How important is knowing some computer science for mathematicians? How much do you use computers in your research? Which fields are most related to mathematics?"

3 Mark Behrens was a professor at MIT and is currently a professor at the Department of Mathematics in the University of Notre Dame. His research focuses on the computational aspects of algebraic topology.

4 Clark Barwick was a professor at MIT and currently a professor at the University of Edinburgh. His research focuses on homotopy theory, algebraic K-theory, and higher category theory.

Professor Strang

I think it's good to be able to do some things with a computer. To be able to do experiments with a computer and to think in that way—which I didn't really get. So I'm in linear algebra, but I'm not really in computational linear algebra. That's not my thing. I'm the MathWorks professor; Mathworks focuses on computation very successfully. Professor Edelman has created the language Julia which is amazing that he could create a new and super fast language. And then Python came out for engineering. And I'm sure statistics, which uses the language R—but that's an older one—I'm sure statistics has got a future super language coming up. Somebody will do that. It's a world in which, when there's an open problem, a space to be filled, somebody jumps in, and sees a way to do it, [and does it].

"There's a lot of competition or pressure from fellow students here that it's [all the more] important to have some confidence in yourself that if you put in some work you can do well."

Professor Guth

I don't really use any [technology or computer science]. I'm not that good with computers, and I don't like spending time on them that much; that's a personal thing. I guess I use search engines to search for papers on different stuff—that is actually an amazing computer tool—and I think it's really somewhat better than the old system with very thick books and references to other papers. I wonder sometimes about using computers to, for example, look for examples for some of these problems. Like I've wondered about using computers to try to find interesting examples for the Kakeya problem. I haven't ever done it because I don't have the computer know-how, but I'm curious about it. And I don't know very much about this stuff but there's a lot of stuff in the news about machine learning recently. So I'm curious about how much machine learning could tell us about the problems that I've worked on and been stuck on.

Professor Hahn

I had a lot of friends who did computer science. I picked up some programming. I used to do some things over IAP, some competitions like Pokerbots that kept my coding skills sharp. I keep an academic interest in computer science. I like to read some literature in some parts of that, but I didn't take too much formal computer science as an undergraduate personally.

[The importance of computer science] really depends on what you want to do. I think that's a great and really valuable career route you want to keep open, even if you want to be a pure mathematician, if you change your mind later. But I think I was lucky to not need that. I tried to be conscious of keeping my coding skills, and general knowledge (of computer science), up a little bit in part to have an alternative career option depending on whether this algebraic topology thing worked out. But for me it seems it's worked out pretty well.

CONCLUSION

From our interview of the mathematics departments' Professor Gilbert Strang, Professor Larry Guth, and Professor Jeremy Hahn, we hear their takes on mathematical research, both their own and for students at MIT. We have seen that industry is indeed an option for mathematicians, but becoming a research mathematician is also a possibility, especially for pure mathematicians. However, it is indeed more competitive.

Nonetheless, the professors encourage people to pursue their interests, regardless of it being purely mathematics, some fields in relation to mathematics, or something else entirely. For mathematics students, Professors Strang, Guth, and Hahn recommended that students explore their strengths, be ready for setbacks, and focus on themselves, respectively.

With regards to looking for UROPs at MIT, professors highlighted looking for students who have taken higher-level math classes. The professor also noted that computational skills are an advantage. However, while many research mathematicians use technology in their research, a significant amount of research is still done without it.

We are very grateful to the professors for allowing us to interview them and for sharing their priceless experiences and advice. For the students reading this, we hope that this article has given you more insight into what mathematical research at MIT is like and answered commonly asked questions about UROPs, research, and more!

REFERENCES

Barwick, C. (n.d.). CLARKE BARWICK. Retrieved from https://www.maths.ed.ac.uk/~cbarwick/

Behrens, M. (n.d.). Home Page for Mark Behrens. Retrieved from https://www3.nd.edu/~mbehren1/

California Institute of Technology. (2023). California Institute of Technology: The Division of Physics, Mathematics, and Astronomy. Nets H. Katz. Retrieved from https://pma.caltech.edu/people/nets-h-katz

Dvir, Z. (n.d.). Zeev Dvir's Homepage. Retrieved from https://pma.caltech.edu/people/nets-h-katz

Hahn, J. (n.d.). Jeremy Hahn. Retrieved from https://web.mit.edu/~jhahn01/www/

Massachusetts Institute of Technology. (n.d.). MIT Mathematics. Larry Guth. Retrieved from https://math.mit.edu/directory/profile?pid=1461

Strang, G. (n.d.). Gilbert Strang. Retrieved from https://math.mit.edu/~gs/it's imt

We chase the miracles of science to *improve people's lives*

Trainee Lucille Pernot, R&D plateform, Vitry-sur-Seine, France - © Vincent Fournier

Sanofi is dedicated to transforming the practice of medicine by working to turn the impossible into the possible. We provide potentially life-changing treatment options and life-saving vaccine protection to millions of people globally, while putting sustainability and social responsibility at the center of our ambitions.

Work with us! Learn more about our partnering interests, collaborations, award programs and careers at:

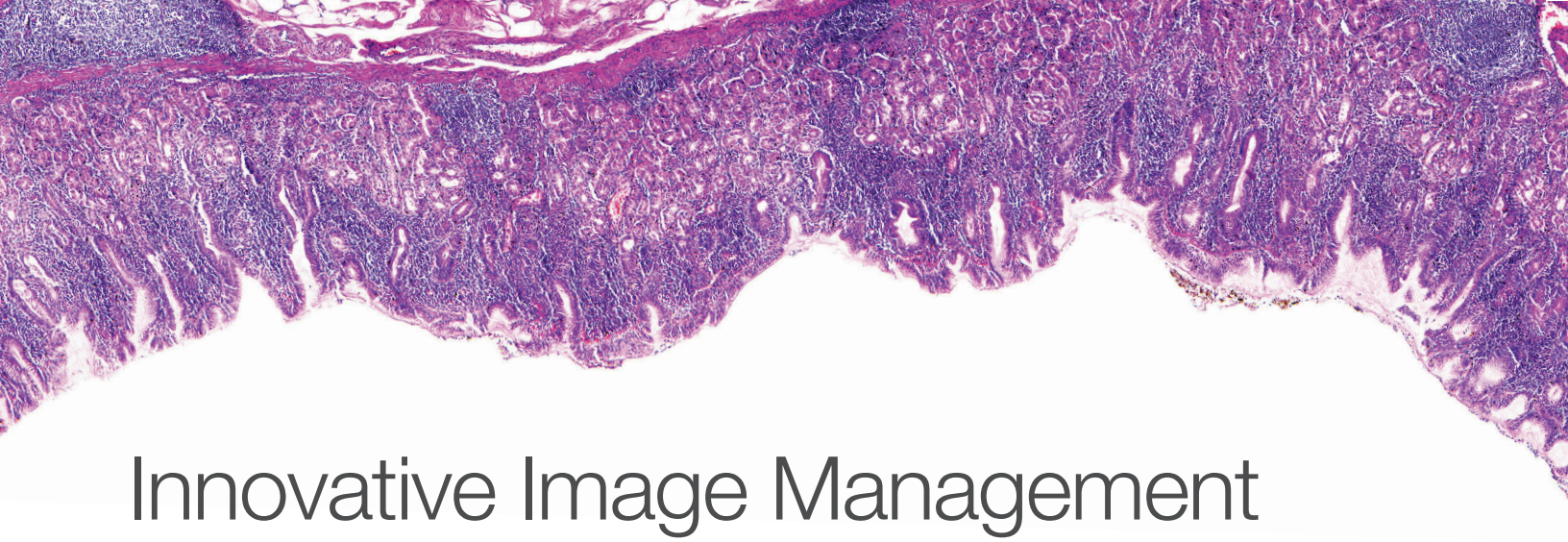
https://www.sanofi.com/en/science-and-innovation

https://www.sanofi.com/en/science-and-innovation/collaboration-and-awards

https://www.sanofi.us/en/careers/

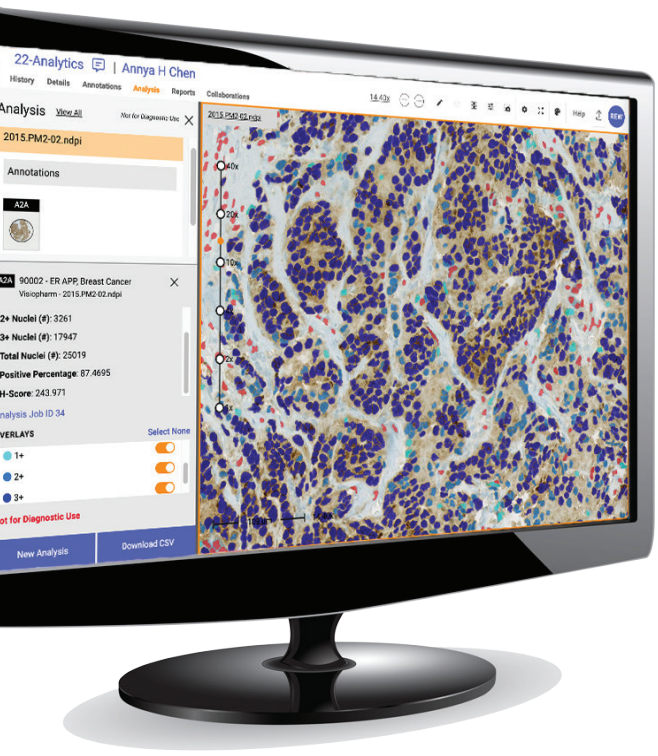
Scan the QRCode below to learn more!

sanofi



Innovative Image Management
For Digital Pathology

Corista's DP3 delivers the industry's most extensive array of tools for digital pathology.



- Bring the power of AI and industry-leading analytics directly into your patient's file for analysis, review and storage.
- Instantly view high-definition images anytime from anywhere in the world.
- Experience next generation digital pathology with a refined user experience and seamless integrations.
- Scale to handle the multiple sites, physicians and scanners of your organization optimizing the workflows for all.

This is Integrated Pathology

corista

CORISTA | Call 978.287.6188 or visit us at corista.com

MURJ

UROP Summaries

MobilePath: A Mobile Server Platform for Health and Behavior Monitoring

Sabrina Do[1], Valerie Kwek[2], Somaia Saba[2], Ariadne Dulchinos[2], Richard Fletcher[3]

1 Student Contributor, Department of Computer Science and Molecular Biology, MIT, Cambridge MA 02139

2 Student Contributor, Department of Computer Science, MIT, Cambridge MA 02139

3 Principal Investigator, Department of Mechanical Engineering, MIT, Cambridge MA 02139

Background

One of the longest-standing challenges in medicine is the ability to monitor and predict human behavior. Applications of behavior medicine include difficult problems such as detecting drug addiction relapse, monitoring patient recovery after surgery, or helping to predict suicide. Over the past decade, the widespread usage of mobile phones has created new opportunities for monitoring tools in the field of behavioral medicine (Insel, 2017; Melcher et al., 2020). Specifically, mobile phones allow for the passive, minimal interference, more precise, objective in-situ collection of data, specifically of one's activity and daily behavior (e.g. sleep, exercise) (Onnela, 2021). The phenotyping problem, or the inability to precisely specify phenotypes is solved by smartphones – the use of this digital data to quantify individual-level human phenotype is known as digital phenotyping (Onnela, 2021; Onnela, 2016; Insel, 2017). One's daily behavior can give doctors clinical insight remotely into the intensity and progression of a patient's disease, such as mental health or the recovery of cancer patients; this objective data can be incredibly significant for fields such as psychiatry as it represents a path to measurement-based care rather than care based solely on clinical judgment (Insel, 2017; Jain et al., 2015). For example, researchers have demonstrated that suicidal behavior could be monitored through Google search data (Gunn and Lester, 2013), and accelerometer data can be used to track activity in patients with cardiometabolic diseases (Jain et al., 2015). This can be incredibly important as doctors can continue to monitor a patient's disease after their visit, especially critical in global health screenings. Furthermore, the need for an asynchronous telehealth solution that can increase access to care has been increasingly critical, as revealed during the COVID-19 pandemic (Melcher et al., 2020).

Results

Overview of MobilePath

Despite this increasing use of mobile technology, a server platform for digital phenotyping designed to visualize and monitor health behavior for both physicians and clinical researchers does not exist. To address this need, our group has developed a data collection and assessment digital phenotyping platform, called MobilePath, which also includes an Android mobile application and server to store health and behavior measurements from mobile devices. MobilePath is an easy-to-use data collection and assessment platform designed for medical researchers and clinicians to better monitor behavior and be able to detect specific patterns in behavior that are associated with adverse behavior, such as substance use relapse, or detect specific psychopathology, such as depression.

Framework

The framework contains features for clinicians and patients to view medical measurements. Specifically, we have developed abstract visualizations for multiple health variables such as accelerometer periodic, accelerometer events, light sensor data, gyroscope orientation, circadian rhythm, phone call and SMS/texting activity, mobility data, GPS movement, GPS log mobility, and phone usage. Figure 1 displays the patient and clinician user interface, specifically the measurements dashboard, and includes an example plot of the patient's periodic accelerometer data collected from the user's mobile device with the Android application. Users are asked to input a date range over which data is collected, and MobilePath visualizes the data from that time period. An example plot graphing the number of times a patient sent and received SMS messages over time is depicted in Figure 2. As individuals with social anxiety or loneliness have been found to use smartphones differently, such as having fewer SMS events and incoming calls, this data on SMS events can be useful for clinicians to identify social anxiety (Gao et al., 2016). As shown in Figure 3, MobilePath also includes a visualization page for researchers. This visualization converts some features, specifically light sensor, GPS, log GPS, and battery life measurements, into a colored square. Specifically, all data features are scaled and normalized – either by probability (0 to 1) or by z-score, where z=0 is mapped to the value 0.5. This scaled and normalized value is then converted to an

Fig. 1. MobilePath's patient website user interface of the measurement dashboard. The measurement dashboard includes all the data collected over a specific inputted date range of that specific patient. A plot of a user's periodic accelerometer data is displayed and graphed on MobilePath from mobile sensor data.

RGB value from white (0, 0, 0) to blue (0, 0, 255). This visualization can be used for machine learning analyses by clinical researchers using MobilePath.

Security and Privacy

To avoid the typical pitfalls and risks of mobile health apps, which primarily concern personal privacy, MobilePath was designed with privacy in mind and operates only with the full consent of the patient (Onnela, 2016; Insel, 2017; Kröger, 2022). Multiple measures, which are explained to each participant at the time of the consent, are taken by MobilePath to protect patient privacy. All data sent from the phone to the server is de-identified and designated with a userid code. Additionally, microphone/sounds and camera data are not collected. While GPS location data is collected, the GPS coordinate is blurred by adding a random 100-mile offset, so that the participant's exact location cannot be ascertained. The server and analysis only keeps track of the relative movement of the phone, not the absolute coordinates. The participants' phone calls and SMS events are logged, but the phone numbers are scrambled (hashed) with a random key so that the phone numbers cannot be ascertained. The only information collected is the time, duration, and scrambled phone number of each call, and the time and scrambled phone number of each SMS, both sent and received. The server does not store any phone numbers; only how many unique numbers were called or received, which is used to inform a measure of social connectedness, are kept track of. In terms of data security, the data is stored on a password-protected server, and all communication between

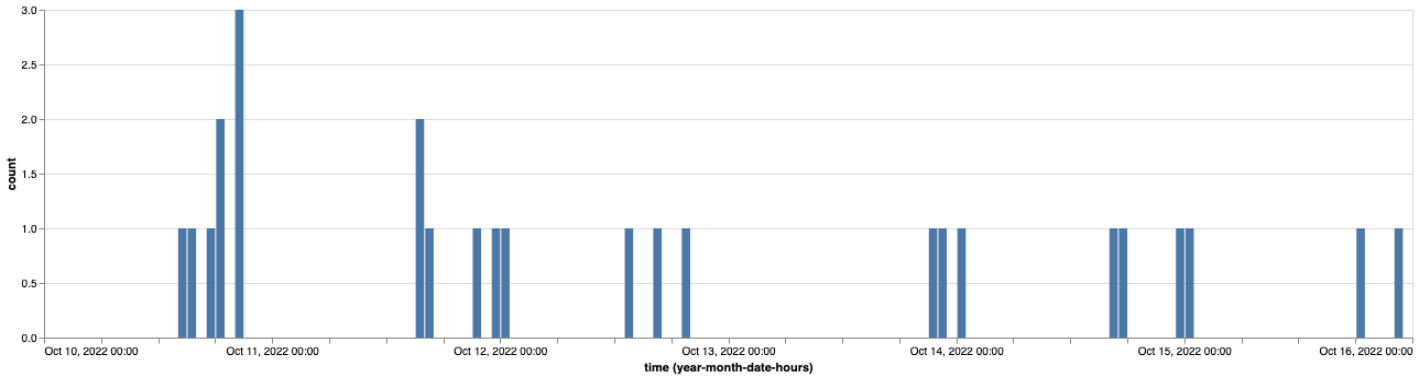


Fig. 2. A plot of a user's incoming and outgoing SMS events each hour displayed on MobilePath based on data collected from mobile sensor data from MobilePath's Android application.

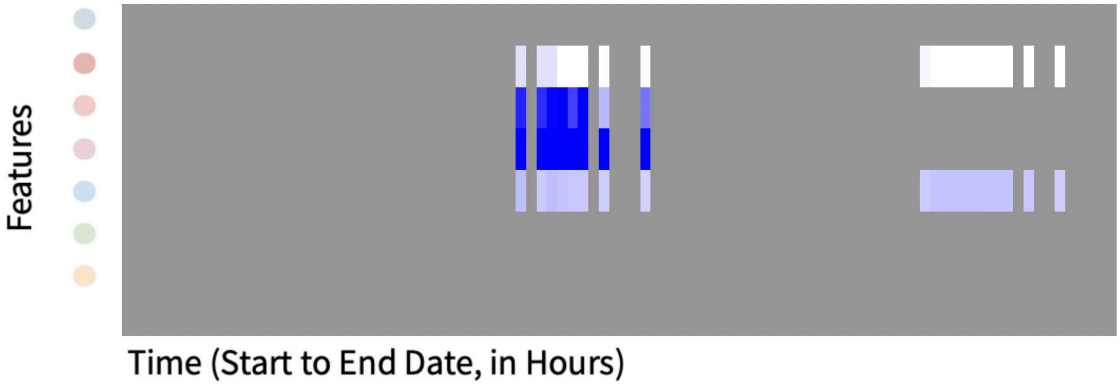


Fig. 3. A plot visualizing a patient's data to be used by clients for machine learning analyses. Features, specifically light sensor, GPS measurements, log GPS measurements, and battery life, are scaled and normalized into a probability or z-score value. This scaled and normalized value is then converted to an RGB value where a probability of 0 is mapped to an RGB value of (0, 0, 0), and a probability of 1 is mapped to an RGB value of (0, 0, 255).

the participant's smartphone and MobilePath's central server is encrypted with an SSL certificate.

Conclusion

MobilePath ultimately serves as a digital phenotyping platform that facilitates the monitoring and prediction of human behavior to improve clinical care and research. Future steps for MobilePath include conducting some pilot clinical studies later in 2023 and using the abstract visualizations along with deep learning algorithms to predict a user's digital phenotype and detect specific behaviors, such as a user's sleep habits, indicators of stress, depression, or anxiety, and scrolling/texting.

References

Gao Y, Li A, Zhu T, Liu X, Liu X. (2016). How smartphone usage correlates with social anxiety and loneliness. PeerJ. 4:e2197. https://doi.org/10.7717/peerj.2197

Gunn J, Lester D. (2013). Using google searches on the internet to monitor suicidal behavior. Disorders, 148(2-3), 411- 412. Retrieved on March, 12, 2023 from https://doi.org/10.1016/j.jad.2012.11.004.

Insel, TR. (2017). Digital Phenotyping: Technology for a New Science of Behavior. JAMA, 318(13), 1215-1216. Retrieved on June 7, 2021 from https://doi.org/10.1001/jama.2017.11295

Jain SH, Powers BW, Hawkins JB, Brownstein JS. (2015). The digital phenotype. Nat Biotechnol, 33(5), 462-463. Retrieved on March 12, 2023 from https://doi.org/10.1038/nbt.3223

Kröger, Jacob & Gellrich, Leon & Brause, Saba & Pape, Sebastian & Ullrich, Stefan. (2022). Personal information inference from voice recordings: User awareness and privacy concerns. Proceedings on Privacy Enhancing Technologies, (1), 6-27. Retrieved on March 12, 2023 from https://doi.org/10.2478/popets-2022-0002

Onnela JP, Rauch S. (2016). Harnessing Smartphone-Based Digital Phenotyping to Enhance Behavioral and Mental Health. Neuropsychopharmacol, 41, 1691-1696. Retrieved on March 12, 2023 from https://doi.org/10.1038/npp.2016.7

Onnela, JP. (2021). Opportunities and challenges in the collection and analysis of digital phenotyping data. Neuropsychopharmacol, 46, 45-54. Retrieved on March 12, 2023 from https://doi.org/10.1038/s41386-020-0771-3

Melcher J, Hays R, Torous J. (2020). Digital phenotyping for mental health of college students: a clinical review. Evid Based Ment Health, 23(4):161-166, Retrieved on Nov 30, 2022 from https://doi.org/10.1136/ebmental-2020-300180.

Electrowetting Valve Manufacturing and Testing for Electrospray Thruster Production

Nigel Barnett[1], Matthew Corrado[2], Paulo Lozano[3]

1 Student Contributor, Department of Aeronautics and Astronautics, MIT, Cambridge MA 02139

2 Supervisor, Department of Aerospace Engineering, MIT, Cambridge MA 02139

3 Principal Investigator, Department of Aerospace Engineering, MIT, Cambridge MA 02139

Nanosatellites are cheap to produce and launch; however, a lack of control after being released into orbit means nanosatellites typically have shorter lifespans than their larger counterparts. The MIT Space Propulsion Laboratory (SPL) has been developing modular electrospray thrusters that satisfy this need for small- scale satellite propulsion. Electrospray thrusters (Fig. 1) are compact engines that incorporate micro-electromechanical systems (MEMS) to combine electric charge and charged fuel to create thrust. Such thrusters could provide control to operators on the ground, allowing them to maneuver satellites away from debris,

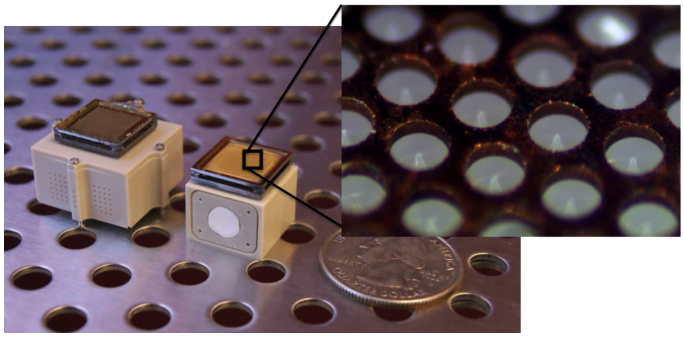


Fig. 1. Electrospray thrusters and emitter grid close-up (Krejci, 2016).

raise satellite orbits to correct for atmospheric drag in low Earth orbit, and perform more complex missions.

The small scale and high efficiency of electrospray propulsion mean that mission planners can add modular thrusters to extend the mission life of a satellite (Greene, 2008). However, while electrospray propulsion has many benefits, the small size of an electrospray thruster presents challenges for manufacturing. Components like emitter arrays and fuel valves must be manufactured precisely, requiring complex microfabrication processes similar to computer chips. This research project focuses on how the SPL can reliably produce electrically actuated valves for electrospray thruster propellant control. These electrowetting valves initially separate an electrospray thruster's fuel tank and emitter (nozzle) until it needs to be fired. At that point, the thruster systems apply a charge across the valve, changing its hydrophobicity, allowing the fuel to reach the emitter and create thrust.

For the duration of a launch, thruster components can face more than 12 Gs of acceleration, intense vibration, and thermal cycling before reaching the ideal conditions of orbital microgravity (Krejci, 2016). Those forces can cause thruster valves to fail and allow fuel to reach the emitters prematurely, leading to malfunctions and reduced performance. These extreme conditions that the

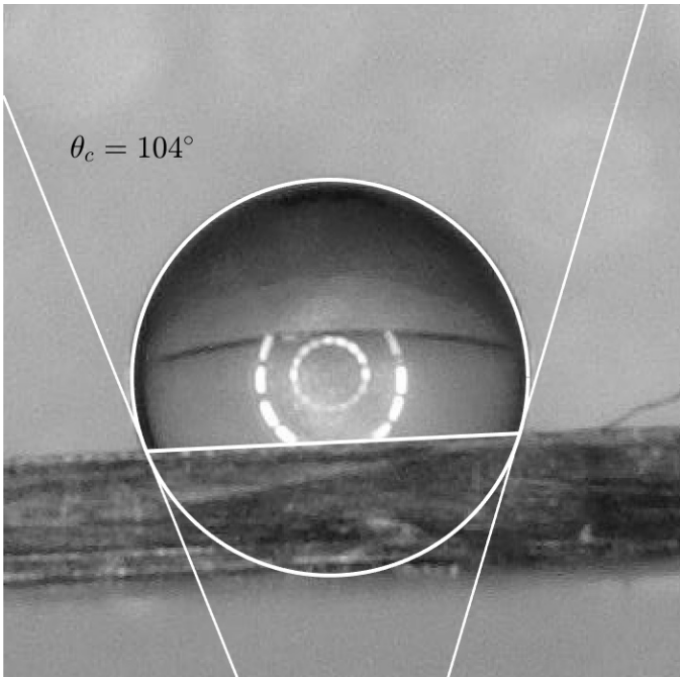


Fig. 2. Contact Angle Measurement of Water with a Sample Valve (See Fig. 3).

	Average Contact Angle (degrees) by Curing Time	
Curing Time	30 Minutes	50 Minutes
Chip Set 1	113	104
Chip Set 2	100	105
Chip Set 3	119	112

Table 1. Merage Measured Contact Angle with Varied Curing Parameters.

thruster experiences and the micro-scale of fabrication make it challenging to manufacture reliable valves.

This project focuses on refining and standardizing the process for manufacturing the electrowetting valves. Each silicon valve substrate is spin coated in a special hydrophobic fluoropolymer. This fluoropolymer, called Fluoropel, is an electrowetting agent that allows the valve to change its interaction with fluids when an electrical field is applied (Kim et al., 2010). We can measure this interaction between the valve and fuel by the contact angle of a drop of fluid. Before the thruster needs to be fired, a large contact angle and high hydrophobicity are optimal so that fuel does not seep through the microscopic valve channels leading to the emit- ter. But when thrust is required, an electrical field can be applied to reduce the contact angle of the valve (Fig. 2) and allow fuel to pass through the valve and power the thruster.

So far, the project has been focused on initial tests of the impact of curing procedures on the contact angle between water and a valve test piece, as shown in figure two. The table below highlights the current results from contact angle testing. Chip set 1 tested the difference in curing times: 30 minutes at 180°C versus 20 minutes at 80°C and then 30 minutes at 180°C. Because of the non-uniformity in the surface finish of the 30-minute valve (Fig. 3), chip set 2 was coated at the same conditions, 3000 rotations per minute (RPM) for five seconds, and cured for the same time as chip set 1. Chip set 2 had a uniform surface finish, so chip set 3 was coated under different spin coating conditions, 5000 RPM for 30 seconds, to test if any trends in contact angle between curing times persisted to the most extreme spin coating conditions.

Given the collected data in Table 1, the current conclusion is that the vendor-recommended curing setting of 30 minutes at 180°C is optimal, even at a more extreme coating of 5000 RPM. However, there is some uncertainty in the significance of the different curing procedures, given that two drops on the same chip could produce up to 12° of difference in measured contact angle.

The next step for the project is to follow a test matrix to characterize the contact angle based on spin coating parameters and find the optimal spin coating settings for the largest contact angle. The test matrix varies two independent components of the fluoropolymer coating procedure: spin duration and RPM. Therefore, a trend in contact angle should be observed by changing these parameters, and the combination with the largest contact angle will be chosen for future electrowetting valve manufacturing.

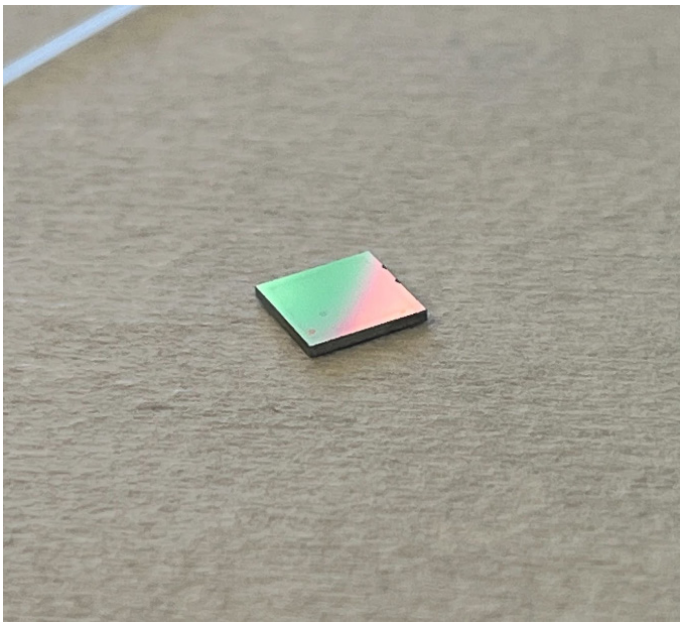


Fig. 3. E7x7mm Valve Sample from Set 1 with Inconsistent Surface Finish.

After determining a suitable procedure for coating the valves, they will be environmentally tested to determine their performance at launch and determine if they are more reliable than previous versions. If they can withstand environmental testing consistently, the project focus will shift to producing more valves for SPL for future thruster production. However, if the valves are unreliable, more work will go into varying the curing settings and spin coating procedure to find a coating procedure that produces a robust valve.

References

Greene, Courtney. 2008. Application of an Electrospray Thruster in a Nanosatellite. Department of Mechanical Engineering. Retrieved on 22 March 2023 from https://dspace.mit.edu/bitstream/handle/1721.1/45839/319618231-MIT.pdf.

Kim, JH., Lee, JH., Mirzaei, A. et al. 2020. Electrowetting-on-dielectric characteristics of ZnO nanorods Nature. Retrieved on 22 March 2023 from https://doi.org/10.1038/s41598-020-71017-7.

Krejci, David. 2016. Electrospray Thruster Engineering. Retrieved on 22 March 2023 from https://spacepropulsion.mit. edu/electrospray-thruster-engineering.thruster experiences and the micro-scale of fabrication make it challenging to manufacture reliable valves.

MURJ Reports

On Symmetric Rank Decompositions of the 3x3 Matrix Multiplication Tensor

Jason Yang[1], Virginia Vassilevska Williams[2]

1 Student Contributor, MIT Class of 2025, Department of Electrical Engineering and Computer Science, Cambridge, MA 02139

2 Supervisor, MIT Computer Science and Artificial Intelligence Laboratory, Cambridge, MA 02139

This paper is an improved attempt to find a rank ≤ 23 mod 2 decomposition of the tensor describing 3 × 3 matrix multiplication. We investigate a much bigger set of possible symmetry restrictions and are more systematic about what kinds of restrictions we test. Our results rule out the existence of a rank ≤ 23 decomposition for many more symmetry restrictions.

1 Introduction

Fast matrix multiplication essentially consists of finding a low-rank decomposition of a $\langle n,k,m\rangle$ *matrix multiplication tensor* $\mathcal{T}^{\langle n,k,m\rangle}$, a $n\times k\times k\times m\times m\times n$ array where

$$\mathcal{T}^{\langle n,k,m\rangle}_{i,j,j,k,k,i}=1\ \forall i,j,k, \text{ and all other elements are 0.}$$

A *R-rank decomposition* over a field $\mathbb{F}$ is a (multi)set of matrix triplets $D=\{(A^{(r)}\in\mathbb{F}^{n\times k},B^{(r)}\in\mathbb{F}^{k\times m},C^{(r)}\in\mathbb{F}^{m\times n})\}_{0\le r<R}$ s.t.

$$\mathcal{T}^{\langle n,k,m\rangle}=\sum_{r=0}^{R-1}A^{(r)}\times B^{(r)}\times C^{(r)} \text{ (equivalently,}$$

$$\mathrm{T}^{\langle n,k,m\rangle}_{a,b,c,d,e,f}=\textstyle\sum_{r=0}^{R-1}A^{(r)}_{a,b}B^{(r)}_{c,d}C^{(r)}_{e,f}\ \ \forall a,b,c,d,e,f),$$

where × denotes the tensor product (equivalent to `numpy.multiply.outer()` from NumPy).

Such a decomposition yields a $O(N^{3\log_{nkm}R})$-time divide-and-conquer algorithm for multiplying two $N\times N$ matrices. [Bläser, 2013] The quantity $3\log_{nkm}R$ is known as the *running time exponent* and is commonly denoted ω. The smallest R such that a R-rank decomposition of $\mathcal{T}^{\langle n,m,k\rangle}$ exists is known as the *rank* of the tensor $\mathcal{T}^{\langle n,m,k\rangle}$.

Let $\langle n,k,m:R\rangle_{\mathbb{F}}$ be short for "a rank-R decomposition of $\mathcal{T}^{\langle n,k,m\rangle}$ over field $\mathbb{F}$". [Strassen, 1969] proved that the trivial algorithm for matrix multiplication is not optimal, by finding a divide-and-conquer algorithm that can be encoded as $\langle 2,2,2:7\rangle_{\mathbb{Z}}$ [1] ($\omega\approx 2.807$). The next improvement came from [Pan, 1978] with $\langle 70,70,70:143640\rangle_{\mathbb{Z}}$ ($\omega\approx 2.795$). Every record algorithm afterward has used asymptotic inequalities to nonconstructively prove that certain running time exponents can be obtained (or approached arbitrarily closely), instead of directly finding decompositions; [Bläser, 2013] the current record achieves a running time of $O(N^{2.37188})$. [Duan et al., 2022]

While these asymptotic techniques allow such algorithms achieve significantly lower asymptotic running time exponents than any known decompositions of small tensors, they also render such

[1] By slight abuse of notation, what we mean here is that the decomposition is over $\mathbb{Q}$ but only uses coefficients in $\mathbb{Z}$.

algorithms too slow for any matrices encountered in practice; thus, there is still much demand for small, explicitly constructed decompositions. Significant progress has been made in this area in recent years, especially $\langle 3,3,6:40\rangle_{\mathbb{Q}}$ ($\omega \approx 2.774$) [Smirnov, 2013] and $\langle 5,5,5:97\rangle_{\mathbb{Z}}$ ($\omega \approx 2.842$) [Kauers & Moosbauer, 2022], and additional results are catalogued at `https://fmm.univ-lille.fr/`. [2]

We focus on $\mathcal{T}^{\langle 3,3,3\rangle}$, since it is the smallest tensor of the form $\mathcal{T}^{\langle n,n,n\rangle}$ whose rank is still unknown; the best-known upper bound is 23 ($\omega \approx 2.854$), first discovered by [Laderman, 1976], which has not been beaten in 40+ years. Inspired by [Ballard et al., 2018], we enforce symmetry restrictions on decompositions to significantly reduce search space, which also motivates us to focus on tensors $\mathcal{T}^{\langle n,n,n\rangle}$. Finally, we search over $\mathbb{Z}_2$ to further reduce search space.

Our objective is to find (or rule out the existence of) a rank ≤ 23 decomposition of $\mathcal{T}^{\langle 3,3,3\rangle}$ over $\mathbb{Z}_2$, under one of several different symmetry restrictions. We find eight decompositions of rank 23 under four different symmetry restrictions (shown in Section 5) and rule out the existence of such decompositions for many more symmetry restrictions, although we still have not found a rank 22 decomposition.

1.1 Notation

We will only consider tensors $\mathcal{T}^{\langle n\rangle} := \mathcal{T}^{\langle n,n,n\rangle}$.

A decomposition $D = \{(A^{(i)}, B^{(i)}, C^{(i)})\}_i$ is a (multi)set of matrix triplets. Borrowing notation from [Kolda & Bader, 2009], $[[D]]$ denotes $\sum_{(A,B,C)\in D} A \times B \times C$, which we call the "tensor valuation" of D. If D is a nested set of matrix triplets, $[[D]]$ denotes the tensor valuation of the flattened version of D.

For a function f that transforms a matrix triplet into another matrix triplet, $f(D)$ is short for $\{f(d) : d \in D\}$.

For a set of functions $F = \{f_i\}_i$, $\langle f_0, f_1, \ldots\rangle$ or $\langle F\rangle$ denotes the generation of F, i.e. the set (group) of all functions that are compositions of finite (possibly empty) sequences of f_i and f_i^{-1}. For a matrix triplet x, $\langle F\rangle(x) = \{f(x)\}_{f\in\langle F\rangle}$ is the orbit of x under $\langle F\rangle$.

$\|$ denotes multiset sum (list concatenation but with multisets). For a (multi)set of matrix triplets X, $\langle F\rangle(X) := \|_{x\in X}\langle F\rangle(x)$. For Q as a set of sets of matrix triplets, $^{\|}Q$ denotes $\|_{q\in Q} q$, i.e. the concatenation of the elements of Q.

For a string/sequence-like object S, $S_{[a:b]}$ denotes the subsequence $S_a, \ldots S_{b-1}$, and $S_{[:b]}$ is short for $S_{[0:b]}$.

Unless stated otherwise, the field $\mathbb{F}$ we are working over will be $\mathbb{Z}_2$.

2 Symmetry and Mod 2 Constraints

We will consider three kinds of transformations on matrix triplets (A, B, C): [De Groote, 1978]

- cycle: $\triangle((A,B,C)) = (B,C,A)$;
- transpose: $\top((A,B,C)) = (C^\top, B^\top, A^\top)$;
- trace ("sandwich"): $\phi_{X,Y,Z}((A,B,C)) = (XAY^{-1}, YBZ^{-1}, ZCX^{-1})$ for invertible $X, Y, Z \in \mathbb{F}^{n\times n}$; [3]

It can be checked manually that for each $f \in \{\triangle, \top, \phi_{X,Y,Z}\}$, if a decomposition D satisfies $[[D]] = \mathcal{T}^{\langle n\rangle}$, then $[[f(D)]] = \mathcal{T}^{\langle n\rangle}$. [4] Because of this property, we say that the tensor $\mathcal{T}^{\langle n\rangle}$ is symmetric under $\triangle$, $\top$, and $\phi_{X,Y,Z}$. By composition, $\mathcal{T}^{\langle n\rangle}$ is symmetric under any function in the group $\mathbf{\Gamma} := \langle \triangle, \top, \phi_{X,Y,Z}\rangle$.

[2] Interestingly, however, the current best-known running time exponent from a small tensor decomposition still seems to be $\langle 44,44,44:36133\rangle_{\mathbb{Q}}$ ($\omega \approx 2.7734$). [Pan, 1982]

[3] The P_σ function we used in our previous paper, for a permutation $\sigma \in S_n$, is equivalent to $\phi_{S,S,S}$ where S is the corresponding permutation matrix of σ

[4] The symmetries under $\triangle$ and $\top$ can be shown via the identities $\mathcal{T}^{\langle n\rangle}_{a,b,c,d,e,f} = \mathcal{T}^{\langle n\rangle}_{c,d,e,f,a,b}$ and $\mathcal{T}^{\langle n\rangle}_{a,b,c,d,e,f} = \mathcal{T}^{\langle n\rangle}_{f,e,c,d,b,a}$. The symmetry under $\phi_{X,Y,Z}$ can be shown by the fact that for arbitrary matrices P, Q, R, the "dot product" of $\mathcal{T}^{\langle n\rangle}$ with $P\times Q\times R$, $\sum_{a,b,c,d,e,f}\mathcal{T}^{\langle n\rangle}_{a,b,c,d,e,f}P_{a,b}Q_{c,d}R_{e,f}$, equals trace$(PQR)$, which equals trace$((XPY^{-1})(YQZ^{-1})(ZRX^{-1}))$ via algebraic properties of the trace function.

2.1 Motivation for symmetry restrictions

Because $\mathcal{T}^{\langle n\rangle}$ is symmetric under each $f \in \mathbf{\Gamma}$, we might guess that there exists some decomposition D of $\mathcal{T}^{\langle n\rangle}$ that happens to satisfy $D = f(D)$ for some f; if this is true, we say that the *decomposition* D is symmetric under f. Such a decomposition would be much easier to find than an arbitrary decomposition of $\mathcal{T}^{\langle n\rangle}$, since there are usually far fewer symmetric decompositions than arbitrary decompositions, up to any given rank.

Furthermore, most explicitly constructed low-rank decompositions of $\mathcal{T}^{\langle n\rangle}$ happen to have some kind of symmetry. [Strassen, 1969] corresponds to $\mathcal{T}^{\langle 2\rangle} = \left\langle \triangle, \phi_{F=\left[\begin{smallmatrix}0&1\\1&0\end{smallmatrix}\right],F,F} \right\rangle (\{(I_2, I_2, I_2)\},$ $([\begin{smallmatrix}0&0\\1&1\end{smallmatrix}], [\begin{smallmatrix}1&0\\0&0\end{smallmatrix}], [\begin{smallmatrix}0&1\\0&-1\end{smallmatrix}])$ over $\mathbb{Z}$, and several rank-23 decompositions of $\mathcal{T}^{\langle 3\rangle}$, including the first to be discovered [Laderman, 1976], have nontrivial symmetry. [Ballard et al., 2018]

2.2 Motivation for working over mod 2

The most sought-after low-rank decompositions of $\mathcal{T}^{\langle n\rangle}$ are those over $\mathbb{Z}$, as the algorithms they yield have lower constant factors and better numerical stability, which are important in practice. Furthermore, most explicit low-rank decompositions of certain $\mathcal{T}^{\langle n\rangle}$ (and more general $\mathcal{T}^{\langle n,k,m\rangle}$) that have been found so far happen to be over $\mathbb{Z}$. [Strassen, 1969] [Laderman, 1976] [Pan, 1978] [Ballard et al., 2018]

Such decompositions are extremely difficult to find. Luckily, using properties of modular arithmetic, any decomposition of $\mathcal{T}^{\langle n\rangle}$ over $\mathbb{Z}$ must also be a decomposition over $\mathbb{Z}_2$, so searching mod 2 can rule out many potential decompositions over $\mathbb{Z}$ while vastly reducing search space. Interestingly, Strassen's algorithm may have been obtained by first solving a system of equations mod 2, then extending it to $\mathbb{Z}$. [Landsberg, 2019]

Disclaimer: searching over mod 2 cannot determine anything about the existence of a non-integer decomposition, or a decomposition symmetric over some function involving non-integers, ex. $\phi_{X=\frac{1}{2}\left[\begin{smallmatrix}-1&\sqrt{3}\\-\sqrt{3}&-1\end{smallmatrix}\right],X,X}$.

Furthermore, if a decomposition symmetric under some functions $f \in \mathbf{\Gamma}$ is found over mod 2, care must be taken when extending it to $\mathbb{Z}$, as there may be multiple functions that are equivalent to f mod 2 but different over $\mathbb{Z}$, and/or the orbits of matrix triplets generated by the symmetry group of the decomposition may have different sizes in mod 2 vs. in $\mathbb{Z}$.

3 Search problem and algorithm

For some integers n, R and some set $F \subseteq \mathbf{\Gamma}$, we want to find some decomposition D of rank $\leq R$ s.t. $[[D]] = \mathcal{T}^{\langle n\rangle}$ and $D = f(D)$ for all $f \in F$.

These constraints on D immediately imply $\langle F\rangle(d) \subseteq D \ \forall d \in D$; thus, D must be a concatenation of orbits of individual matrix triplets under $\langle F\rangle$. Our search problem is then equivalent to finding a set of orbits $Q \subseteq \Theta := \left\{\langle F\rangle(r) : r \in (\mathbb{F}^{n\times n})^3\right\}$ s.t. $\left[\left[{}^{\|}Q\right]\right] = \sum_{q\in Q}[[q]] = \mathcal{T}^{\langle n\rangle}$ and $|{}^{\|}Q| = \sum_{q\in Q}|q| \leq R$; then $D =^{\|} Q$. Since we are working over mod 2, WLOG we can assume Q contains distinct orbits.

3.1 Nullspace compression

It can be shown that for any decomposition D and function $f \in \{\triangle, \top, \phi_{X,Y,Z}\}$, applying f on D transforms $[[D]]$ linearly: there is a matrix M_f s.t. $\overrightarrow{[[f(D)]]} = M_f\overrightarrow{[[D]]}$ for any decompostion D, where $\overrightarrow{T}$ denotes vectorization of a tensor T. Since linear transformations are closed under composition, this is also true for any $f \in \mathbf{\Gamma}$.

Thus, for any $T = [[q]]$ for an orbit $q \in \Theta$, we have $\overrightarrow{T} = M_f\overrightarrow{T} \ \forall f \in F$ $\Rightarrow \overrightarrow{T} \in \text{nullspace}\left(W := \left[\begin{array}{c}\vdots\\ \hline M_f - I\\ \hline \vdots\end{array}\right]_{f\in F}\right)$. Performing row reduction on W yields a basis matrix B and a list of indices β s.t. for any $\overrightarrow{T} \in \text{nullspace}(W)$, $\overrightarrow{T} = B\overrightarrow{v}$

for unique $\vec{v} = \begin{bmatrix} \vdots \\ (\vec{T})_{\beta_i} \\ \vdots \end{bmatrix}_i$; we call $\vec{v}$ the *compression* of $\vec{T} \in \text{nullspace}(W)$ and denote it as $\vec{\vec{T}}$. We can then replace our search condition $\sum_{q\in Q}[[q]] = \mathcal{T}^{\langle n\rangle}$ with $\sum_{q\in Q}\overrightarrow{[[q]]} = \overrightarrow{\mathcal{T}^{\langle n\rangle}}$ and deal with compressed tensors instead of tensors themselves.

Letting $\mathrm{E}_{i,j}$ denote the $n\times n$ matrix with 1 at (i,j) and 0 everywhere else, $\{\mathrm{E}_{a,b}\times\mathrm{E}_{c,d}\times\mathrm{E}_{e,f}\}_{a,b,c,d,e,f}$ forms a basis for all 6-dimensional tensors of side length n, so M_f can be constructed by column-wise concatenating tensors $[[f((\mathrm{E}_{a,b},\mathrm{E}_{c,d},\mathrm{E}_{e,f}))]]$.

3.2 Prefix matching

Before proceeding with the search algorithm, we choose a small number b and see if there exists a solution Q s.t. $\overrightarrow{[[{}^{\|}Q]]}$ and $\overrightarrow{\mathcal{T}^{\langle n\rangle}}$ match at the first b bits.

We can solve this decision problem with dynamic programming. To simplify the problem, we allow Q to have duplicate orbits.

Define $\Lambda(r)$ denote the set of all possible b-bit prefixes of all $Q \subseteq \Theta$ s.t. $|Q| \le r$; then we have set $\Lambda(0) = \{\overrightarrow{0}\}$ and $\Lambda(r) = \{v \oplus \overrightarrow{[[q]]}_{[:b]} : q \in \Theta, |q| \le r, v \in \Lambda(r-|q|)\}$. The idea is to fix one of the elements of a Q with an orbit of rank r and recurse on the rest of the orbits.

To simplify the dynamic programming, we allow Q to have duplicate orbits, so $\Lambda(r)$ may contain unnecessary bit-vectors, but this does not affect correctness and does not significantly affect running time. The asymptotic time complexity of this procedure is $O(2^b\cdot|\Theta|\cdot R)$, since each $\Lambda(r)$ must have size $\le 2^b$, and at most $|\Theta|$ many q and 2^b many v for each q are considered while constructing $\Lambda(r)$. However, since most orbits in Θ have relatively large rank, the actual running time is usually much lower.

After this procedure, if $\overrightarrow{\mathcal{T}^{\langle n\rangle}}_{[:b]} \notin \Lambda(r)$ for any $0 \le r \le R$, we know there is no solution Q for the original search problem. Due to time and memory constraints, we run this dynamic programming for each $b = 1\ldots25$, terminating early if for some b we rule out the possibility of a solution to the search problem.

3.3 Single-orbit optimization

We collect all orbits $q \in \Theta$ of rank $\le R$ and group them by their compressed tensor valuation $\overrightarrow{[[q]]}$. In each group, we choose the orbit with smallest rank, and break ties by smallest *code number*, where the code number of a matrix triplet (A,B,C) is defined as $\$((A,B,C)) := \#(A) + \#(B)2^{(n^2)} + \#(C)2^{(2n^2)}$, with $\#(M) := \sum_{0\le i<n^2} M_{\lfloor \frac{i}{n} \rfloor, i \bmod n} 2^i$. We also remove every orbit with an all-0s compressed tensor valuation.

The remaining orbits are called *canonical*; we will only work with these orbits, since given any tensor decomposition, one can always replace each triplet with its respective representative canonical orbit, and preserve tensor valuation without increasing rank.

3.4 Meet-in-the-middle

We use a *meet-in-the-middle* (MITM) search strategy, which relies on the property $[[Q\|Q']] = [[Q]] + [[Q']]$.

- Create two sets $\breve{S}, \underline{S}$ s.t. $\{\breve{Q}\|\underline{Q} : \breve{Q} \in \breve{S}, \underline{Q} \in \underline{S}\}$ contains all Q we want to search over; [5]
- Store the elements of $\breve{S}$ in a dictionary $\breve{M} = \{\overrightarrow{[[\breve{Q}]]} \to \breve{Q} : \breve{Q} \in \breve{S}\}$ with equal keys resolved by choosing a $\breve{Q}$ with minimal rank;
- Iterate over each $\underline{Q} \in \underline{S}$ and query if $v = \overrightarrow{\mathcal{T}^{\langle n\rangle}} - \overrightarrow{[[\underline{Q}]]}$ is a key in the dictionary: if so, then $Q = M.\text{get}(v)\|\underline{Q}$ is a decomposition of $\mathcal{T}^{\langle n\rangle}$;

3.4.1 Finding a meet-in-the-middle scheme

To find such $\breve{S}, \underline{S}$, we partition the set of all Q we want to search over by what we call the *profile* of an orbit set, defined as $\varrho(Q) = [(\#q \in Q \text{ s.t. } |q| = k)]_{0\le k<K}$,

[5] Unlike in our previous paper, here we only use one pair of sets instead of multiple pairs, to make the MITM simpler.

[6] Note that $\varrho(Q)_0 = 0$ since an orbit cannot have size 0.

where $K = 1 + \max_{q\in\Theta} |q|$; a profile is essentially a histogram of orbit sizes. [6]

Using the fact that $\varrho(Q\|Q') = \varrho(Q) + \varrho(Q')$, where + here denotes element-wise addition, we create two sets of profiles $\breve{P}, \underline{P}$ s.t. $\{\breve{p} + \underline{p} : \breve{p} \in \breve{P}, \underline{p} \in \underline{P}\}$ contains all possible profiles we need to consider; then construct $\breve{S} = \{Q : \varrho(Q) \in \breve{P}\}, \underline{S} = \{Q : \varrho(Q) \in \underline{P}\}$.

Let $H = \varrho(\Theta)$ be the histogram of orbit sizes of the set of all orbits. Since we are working over mod 2, duplicate orbits cancel each other out in tensor valuation, so WLOG we only need to consider profiles p s.t. $p_k \le H_k$ for each k. Let P be the set of all such p with the additional constraint $\sum_k kp_k \le R$ (i.e. total rank is $\le R$): this is the set of all profiles we need to consider, and it can be generated by DFS.

Given a profile p, the expression $c(p) = \prod_{0\le i<K} \binom{H_k}{p_k}$ is the number of sets of orbits with profile p. In a later section we introduce methods to make a query on the dictionary $\breve{M}$ run in practically constant time. Thus, the total cost of our algorithm will be roughly proportional to $|\breve{S}| + |\underline{S}| = \sum_{p\in\breve{P}} c(p) + \sum_{p\in\underline{P}} c(p)$. We call this the "MITM cost" and denote it as $\omega(\breve{P}, \underline{P})$. We seek to minimize it subject to the completeness constraint $P \subseteq \{\breve{p} + \underline{p} : \breve{p} \in \breve{P}, \underline{p} \in \underline{P}\}$ and an arbitrarily chosen memory constraint $|\breve{S}| = \sum_{p\in\breve{P}} c(p) \le m = 5 \cdot 10^8$.

Any optimal solution can be constructed as follows: for each $p \in P$, choose some pair of profile $a^{(p)}, b^{(p)} \in P$ s.t. $a^{(p)} + b^{(p)} = p$; then set $\breve{P} = \{a^{(p)} : p \in P\}, \underline{P} = \{b^{(p)} : p \in P\}$. Essentially, we specify some method of creating each $p \in P$ using a sum-pair, then construct $\breve{P}, \underline{P}$ to contain all necessary profiles to satisfy this specification without including extra unnecessary profiles.

We use hill climbing to choose these pairs and obtain a good heuristic solution:

- Put all profiles of P in some arbitrary ordered list L and initialize $\breve{P} \leftarrow \varnothing, \underline{P} \leftarrow \varnothing$;
- For each p in L, find a pair $a, b \in P$ minimizing $\omega(\breve{P}\cup\{a\}, \underline{P}\cup\{b\})$ and satisfying $\left|\breve{P}\cup\{a\}\right| \le m$, then add a to $\breve{P}$ and b to $\underline{P}$;
- Repeat the above steps 5000 times, each time swapping a random pair of elements in L and keeping this change if it produces a solution with equal or less total cost

3.4.2 Fast queries

After creating the dictionary $\breve{M} := \{\overrightarrow{[[\breve{Q}]]} \to \breve{Q} : \breve{Q} \in \breve{S}\}$, we construct bitsets of all keys of M at specific chunks: specifically, we construct $\breve{\Xi}_w = \{k_{[wW:(w+1)W]} : \text{key } k \text{ in } \breve{M}\}$ for each $0 \le w < \left\lceil \frac{B}{W} \right\rceil$, where B is the length of the compressed tensors, $W = 32$ is a constant, and compressed tensors are padded with 0s as needed.

When querying whether a key k exists in $\breve{M}$, we first query for each w whether $k_{[wW:(w+1)W]}$ is in $\breve{\Xi}_w$ (which is done in constant time); if the answer is yes for all k, then we binary search for k among the keys of $\breve{M}$; otherwise we immediately know k is not in $\breve{M}$.

In the test cases we run, only about $\frac{1}{1000}$ or less of all queries pass these checks for all w, and we usually manage to process 10-50 million queries per second.

3.5 Miscellaneous Optimizations

We use multi-key quicksort to sort the keys of $\breve{M}$ by (inverse) lexicographic order (we order two bit strings a, b of equal length by their last bit, tiebreak by their next last bit, and so on).

Key-value pairs in $\breve{M}$ are stored implicitly. For each profile $p \in \breve{P}$, we enumerate all orbit sets $\breve{Q}$ with profile p via DFS; instead of storing the entire orbit set $\breve{Q}$, we split it into $\kappa \cup \{q\}$, where q is the most recent matrix triplet processed in the DFS. We collect all distinct orbit subsets κ encountered during the DFSs of all profiles into a list $\mathcal{K}$. We associate each orbit $q \in \Theta$ with some arbitrary representative matrix triplet t s.t. $\langle F\rangle(t) = q$, and finally store each orbit set $\breve{Q} = \kappa\cup\{q\}$ as a number $i(2^{3n^2}) + \$(q)$, where $\mathcal{K}_i = \kappa$ and $\$(q)$ denotes the code number of q (defined in an earlier section). As the number of elements in $\mathcal{K}$ is usually a small fraction (usually about $\frac{1}{1000}$) of the number of keys in $\breve{M}$, storing keys this way saves

[7] This entire process is much simpler than is was in our previous paper.

a significant amount of memory. [7]

3.6 Properties of search results

Our algorithm satisfies the following lemma:

Lemma 1. *If there exists a decomposition* $^{\|}\mathcal{Q}$ *of* $\mathcal{T}^{\langle n\rangle}$ *of rank* $\le R$ *for some* $\mathcal{Q} \subseteq \Theta$*, then some (possibly nonequivalent) decomposition of rank* $\le R$ *must be returned by the search algorithm.*

Proof. Let Q be $\mathcal{Q}$ but with each orbit replaced with its corresponding canonical orbit. Then $^{\|}Q$ is a decomposition of $\mathcal{T}^{\langle n\rangle}$ with rank $\le |^{\|}\mathcal{Q}| \le R$, and $Q \subseteq \Theta$.

Let $p = \varrho(Q)$. By the completeness requirement of the MITM scheme, there must exist some $\breve{p} \in \breve{P}, \underline{p} \in \underline{P}$ s.t. $p = \breve{p} + \underline{p}$. Q can then be partitioned into $\breve{Q}, \underline{Q}$ s.t. $\varrho\left(\breve{Q}\right) = \breve{p}$ and $\varrho\left(\underline{Q}\right) = \underline{p}$.

Since $\breve{M}$ must contain the compressed tensor valuations of all orbit sets with profile in $\breve{P}$, $\left[\overrightarrow{\left[\breve{Q}\right]}\right]$ must be a key contained in $\breve{M}$; let $\breve{Q}' = \breve{M}.\text{get}(\left[\overrightarrow{\left[\breve{Q}\right]}\right])$. Since during the construction of $\breve{M}$ we specified that equal keys are resolved by choosing a value with minimal rank, $|^{\|}\breve{Q}'| \le |^{\|}\breve{Q}|$.

Since we iterate over every orbit sets with a profile in $\underline{P}$, at some point we will process $\underline{Q}$. Since $Q = \breve{Q}\|\underline{Q}$, $\left[\overrightarrow{[Q]}\right] = \left[\overrightarrow{\left[\breve{Q}\right]}\right] + \left[\overrightarrow{\left[\underline{Q}\right]}\right] = \overrightarrow{\mathcal{T}^{\langle n\rangle}}$, so querying the key $\overrightarrow{\mathcal{T}^{\langle n\rangle}} - \left[\overrightarrow{\left[\underline{Q}\right]}\right]$ in $\breve{M}$ will yield the value $\breve{Q}'$, and the algorithm will return $Q' = \breve{Q}'\|\underline{Q}$, which satisfies $\left[\left[^{\|}Q'\right]\right] = \mathcal{T}^{\langle n\rangle}$ and $|^{\|}Q'| \le |^{\|}Q| \le R$. □

Taking the contrapositive, if the search algorithm does not yield a decomposition of rank $\le R$, then such a decomposition does not exist.

Since our meet-in-the-middle scheme covers all orbit sets of total rank $\le R$, we can replace R in the lemma with any quantity smaller than R, and the lemma would still hold. This yields the following corollary:

Corollary 1. *Let* R^* *be the smallest possible rank for a solution* Q *to the search problem. If* $R^* \le R$*, then the search algorithm is guaranteed to output a solution with rank* R^**.*

Finally, we note an **important disclaimer: the search algorithm may output a solution with rank** $> R$; this is possible since the set $\left\{\breve{p} + \underline{p} : \breve{p} \in \breve{P}, \underline{p} \in \underline{P}\right\}$ may contain extraneous profiles with rank $> R$. This is not a problem since the algorithm is allowed to output multiple solutions, only one of which has to be optimal.

4 Enumerating Symmetry Subgroups

For all of our search results, we set $n = 3$ (i.e. we searched for decompositions of $\mathcal{T}^{\langle 3\rangle}$).

All symmetry sets F we consider will be of the form $S \cup \{f\}$ for a small subset $S \subseteq \mathbf{\Gamma}$ chosen beforehand and an arbitrary element $f \in \mathbf{\Gamma}$. For this paper we have chosen $S = \varnothing, \{\triangle\}, \{\triangle, \top\}$.

4.1 Conjugacy

Let G be the set of all functions that a given decomposition D is symmetric to. We have that G must be a group, and that conjugation by an arbitrary function that transforms matrix triplets yields $(gGg^{-1})(gD) = gD$, so the existence of a rank $\le R$ decomposition of $\mathcal{T}^{\langle n\rangle}$ symmetric over G implies the existence of rank $\le R$ decomposition over any conjugate of G. Thus, WLOG we only need to consider symmetry groups that are distinct up to conjugacy.

4.2 Optimizations

For a fixed set of functions $S \subseteq \mathbf{\Gamma}$, we want to enumerate all $\langle S \cup \{f\}\rangle$ for $f \in \mathbf{\Gamma}$ that are distinct up to conjugacy. Doing this naively by constructing $\langle S \cup \{f\}\rangle$ for every f and testing conjugacy between pairs of such groups will be slow, as even though $|S \cup \{f\}|$ is small, $|\langle S \cup \{f\}\rangle|$ can be as large as $|\mathbf{\Gamma}| = 28449792$. [8] However, we can use several group-theoretic observations to take advantage of the fact that all generating

[8] In fact, $\mathbf{\Gamma} = \left\langle \triangle,\ \phi_{\left[\begin{smallmatrix}1&0&1\\1&1&0\\0&0&1\end{smallmatrix}\right],I,I} \circ \top \right\rangle$.

sets we deal with are small, and greatly reduce memory and time usage:

- For any group G and set S, $G \supseteq S \Leftrightarrow G \supseteq \langle S \rangle$;
- For any element g and set S, $\left\langle gSg^{-1} \right\rangle = g \langle S \rangle g^{-1}$;

These lemmas directly imply the following conjugacy tests:

- $g \langle S \rangle g^{-1} = \langle T \rangle \Leftrightarrow (gSg^{-1} \subseteq \langle T \rangle$ and $\langle S \rangle \supseteq g^{-1}Tg)$;
- $\langle S \rangle = g \langle S \rangle g^{-1} \Leftrightarrow \langle S \rangle \supseteq gSg^{-1}$ (since $|\langle S \rangle| = |g \langle S \rangle g^{-1}|$);

Furthermore, to detect whether $\langle S \rangle$ and $\langle T \rangle$ are conjugate to each other, we only need to test whether $g \langle S \rangle g^{-1} = \langle T \rangle$ for g that are left coset representatives of $\mathrm{N}_{\mathbf{\Gamma}}(\langle S \rangle)$, the normalizer of $\langle S \rangle$ in $\mathbf{\Gamma}$.

Additionally, we do not need to iterate over every f to find all $\langle S \cup \{f\} \rangle$ up to conjugacy, due to the following identities:

- $\langle S \cup \{f\} \rangle = \left\langle S \cup \{f^k\} \right\rangle$ for k coprime to $\mathrm{ord}(f)$;
- $\langle S \cup \{f\} \rangle = \langle S \cup \{sf\} \rangle = \langle S \cup \{fs\} \rangle$ for $s \in \langle S \rangle$;
- $\eta \langle S \cup \{f\} \rangle \eta^{-1} = \left\langle S \cup \{\eta f \eta^{-1}\} \right\rangle$ for $\eta \in \mathrm{N}_{\mathbf{\Gamma}}(\langle S \rangle)$;

Thus, after processing some f, we can transform f to $f^k, sf, fs, \eta f \eta^{-1}$ for k, s, η satisfying the constraints above, then transform those elements as such and repeat until all reachable elements have been obtained, and skip these elements while enumerating subgroups.

Finally, to store the elements of $\mathbf{\Gamma}$ efficiently, we note that $\top \circ \triangle = \triangle^2 \circ \top$, $\triangle \circ \phi_{A,B,C} = \phi_{B,C,A} \circ \triangle$, and $\top \circ \phi_{A,B,C} = \phi_{(A^{-1})^\top,(C^{-1})^\top,(B^{-1})^\top} \circ \top$, implying $\mathbf{\Gamma} = \left\{\phi_{A,B,C} \circ \triangle^c \circ \top^t : \text{invertible } A, B, C \in {\mathbb{Z}_2}^{3\times 3}\right.$; $0 \le c < 3$; $0 \le t < 2$; thus, each element of $\mathbf{\Gamma}$ can be stored with a $O(1)$-space canonical representation.

5 Solutions

For each $S = \varnothing, \{\triangle\}, \{\triangle, \top\}$, the number of distinct subgroups $\langle S \cup \{f\} \rangle$ up to conjugacy is 60, 73, and 43 respectively. For each F we tested, we initialized R to 23 and ran the dynamic programming for $b = 1 \ldots 25$; if no such b ruled out the possibility of a solution, we then repeatedly decremented R until the MITM cost was at most 500 billion, then ran the search algorithm for the resulting rank R. [9]

Below we list all decompositions of $\mathcal{T}^{\langle 3 \rangle}$ that we managed to find, as well as the conjugacy relations between their symmetry subgroups. We denote jGj^{-1} as $j \star G$.

Several of the symmetries below happen to be conjugate to $G_{\text{Ballard}} = \left\langle \triangle, \phi_{A=\left[\begin{smallmatrix}0&0&1\\1&0&1\\0&1&1\end{smallmatrix}\right],A,A} \right\rangle$, which is the mod 2 version of $\left\langle \triangle, \phi_{\alpha=\left[\begin{smallmatrix}0&0&-1\\1&0&-1\\0&1&-1\end{smallmatrix}\right]\in\mathbb{Z}^{3\times 3},\alpha,\alpha} \right\rangle$, under which a rank-23 decomposition of $\mathcal{T}^{\langle 3 \rangle}$ over $\mathbb{Z}$ was found by [Ballard et al., 2018].

All solutions listed here except the last one have rank 23; the last solution has rank 27.

[9] We used the same computer as last time (2019 MacBook Pro, 2.3 GHz 8-Core Intel Core i9, 16 GB 2667 MHz DDR4), but a different compiler flag that offers more memory: `-Xmx30g`.

$S = \varnothing$:

$\langle F \rangle = \ldots$	conjugacy	$D = \langle F \rangle (E)$, where $E = \ldots$
$\left\langle \phi_{\left[\begin{smallmatrix}1&0&1\\1&1&0\\0&0&1\end{smallmatrix}\right],I,I} \circ \triangle \right\rangle$	$= \left(\phi_{\left[\begin{smallmatrix}0&1&1\\0&0&1\\1&0&1\end{smallmatrix}\right],\left[\begin{smallmatrix}0&1&1\\1&0&0\\1&0&1\end{smallmatrix}\right],\left[\begin{smallmatrix}1&1&0\\0&1&0\\1&0&1\end{smallmatrix}\right]} \circ \top \right) \star G_{\text{Ballard}}$	$\left\{ \left(\left[\begin{smallmatrix}0&1&1\\0&0&0\\0&1&1\end{smallmatrix}\right], \left[\begin{smallmatrix}0&0&0\\0&0&0\\1&1&0\end{smallmatrix}\right], \left[\begin{smallmatrix}1&1&0\\0&0&0\\0&0&0\end{smallmatrix}\right] \right) \right.$ $\left(\left[\begin{smallmatrix}1&0&1\\1&0&1\\0&0&0\end{smallmatrix}\right], \left[\begin{smallmatrix}1&0&0\\1&0&0\\0&0&0\end{smallmatrix}\right], \left[\begin{smallmatrix}1&0&1\\0&0&0\\0&0&0\end{smallmatrix}\right] \right)$ $\left(\left[\begin{smallmatrix}1&0&1\\1&0&1\\0&1&1\end{smallmatrix}\right], \left[\begin{smallmatrix}1&0&0\\0&1&1\\1&1&1\end{smallmatrix}\right], \left[\begin{smallmatrix}1&1&1\\0&0&0\\0&1&0\end{smallmatrix}\right] \right)$ $\left(\left[\begin{smallmatrix}0&1&1\\0&1&1\\0&1&1\end{smallmatrix}\right], \left[\begin{smallmatrix}1&1&1\\0&0&0\\1&1&1\end{smallmatrix}\right], \left[\begin{smallmatrix}0&0&0\\0&0&0\\0&1&0\end{smallmatrix}\right] \right)$ $\left. \left(\left[\begin{smallmatrix}1&0&1\\1&1&1\\0&0&1\end{smallmatrix}\right], \left[\begin{smallmatrix}1&0&0\\0&1&1\\0&0&1\end{smallmatrix}\right], \left[\begin{smallmatrix}1&0&0\\0&1&1\\0&0&1\end{smallmatrix}\right] \right) \right\}$
		$\left\{ \left(\left[\begin{smallmatrix}1&0&0\\1&0&0\\0&0&0\end{smallmatrix}\right], \left[\begin{smallmatrix}1&0&1\\1&0&1\\0&0&0\end{smallmatrix}\right], \left[\begin{smallmatrix}1&0&0\\0&0&0\\0&0&0\end{smallmatrix}\right] \right) \right.$ $\left(\left[\begin{smallmatrix}1&1&1\\1&0&0\\0&1&1\end{smallmatrix}\right], \left[\begin{smallmatrix}0&1&0\\0&1&0\\1&1&1\end{smallmatrix}\right], \left[\begin{smallmatrix}1&0&0\\0&1&0\\0&1&0\end{smallmatrix}\right] \right)$ $\left(\left[\begin{smallmatrix}0&0&0\\0&1&0\\0&1&0\end{smallmatrix}\right], \left[\begin{smallmatrix}1&1&1\\1&1&1\\1&1&1\end{smallmatrix}\right], \left[\begin{smallmatrix}1&1&0\\0&0&0\\0&0&0\end{smallmatrix}\right] \right)$ $\left(\left[\begin{smallmatrix}0&0&0\\0&1&1\\0&1&1\end{smallmatrix}\right], \left[\begin{smallmatrix}0&0&0\\0&0&0\\1&1&1\end{smallmatrix}\right], \left[\begin{smallmatrix}0&1&0\\0&1&0\\0&1&0\end{smallmatrix}\right] \right)$ $\left. \left(\left[\begin{smallmatrix}1&0&1\\1&1&0\\0&0&1\end{smallmatrix}\right], \left[\begin{smallmatrix}1&0&0\\0&1&0\\0&0&1\end{smallmatrix}\right], \left[\begin{smallmatrix}1&0&0\\0&1&0\\0&0&1\end{smallmatrix}\right] \right) \right\}$

$S = \{\triangle\}$:

$\langle F \rangle = \ldots$	conjugacy	$D = \langle F \rangle (E)$, where $E = \ldots$
$\left\langle \triangle, \phi_{\left[\begin{smallmatrix}1&0&1\\1&1&0\\0&0&1\end{smallmatrix}\right],\left[\begin{smallmatrix}1&0&1\\1&1&0\\0&0&1\end{smallmatrix}\right],\left[\begin{smallmatrix}1&0&1\\1&1&0\\0&0&1\end{smallmatrix}\right]} \right\rangle$	$= \left(\phi_{\left[\begin{smallmatrix}1&1&0\\0&1&0\\1&0&1\end{smallmatrix}\right],\left[\begin{smallmatrix}1&1&0\\0&1&0\\1&0&1\end{smallmatrix}\right],\left[\begin{smallmatrix}1&1&0\\0&1&0\\1&0&1\end{smallmatrix}\right]} \circ \top \right) \star G_{\text{Ballard}}$	$\left\{ \left(\left[\begin{smallmatrix}0&0&0\\1&1&1\\1&1&1\end{smallmatrix}\right], \left[\begin{smallmatrix}1&1&0\\1&1&0\\1&1&0\end{smallmatrix}\right], \left[\begin{smallmatrix}0&1&0\\0&0&0\\0&0&0\end{smallmatrix}\right] \right), \right.$ $\left(\left[\begin{smallmatrix}1&0&1\\1&1&0\\0&0&1\end{smallmatrix}\right], \left[\begin{smallmatrix}1&0&1\\1&1&0\\0&0&1\end{smallmatrix}\right], \left[\begin{smallmatrix}1&0&1\\1&1&0\\0&0&1\end{smallmatrix}\right] \right),$ $\left(\left[\begin{smallmatrix}1&0&0\\0&0&0\\0&0&0\end{smallmatrix}\right], \left[\begin{smallmatrix}1&0&0\\0&0&0\\0&0&0\end{smallmatrix}\right], \left[\begin{smallmatrix}1&0&0\\0&0&0\\0&0&0\end{smallmatrix}\right] \right),$ $\left(\left[\begin{smallmatrix}0&1&0\\0&1&0\\0&1&0\end{smallmatrix}\right], \left[\begin{smallmatrix}0&1&0\\0&1&0\\0&1&0\end{smallmatrix}\right], \left[\begin{smallmatrix}0&1&0\\0&1&0\\0&1&0\end{smallmatrix}\right] \right),$ $\left. \left(\left[\begin{smallmatrix}1&0&1\\1&0&1\\0&1&0\end{smallmatrix}\right], \left[\begin{smallmatrix}1&0&1\\1&0&1\\0&1&0\end{smallmatrix}\right], \left[\begin{smallmatrix}1&0&1\\1&0&1\\0&1&0\end{smallmatrix}\right] \right) \right\}$
		$\left\{ \left(\left[\begin{smallmatrix}1&1&0\\1&1&0\\1&1&0\end{smallmatrix}\right], \left[\begin{smallmatrix}0&0&0\\1&1&1\\1&1&1\end{smallmatrix}\right], \left[\begin{smallmatrix}0&1&0\\0&0&0\\0&0&0\end{smallmatrix}\right] \right), \right.$ $\left(\left[\begin{smallmatrix}1&0&1\\1&1&1\\0&0&1\end{smallmatrix}\right], \left[\begin{smallmatrix}1&0&1\\1&1&1\\0&0&1\end{smallmatrix}\right], \left[\begin{smallmatrix}1&0&1\\1&1&1\\0&0&1\end{smallmatrix}\right] \right),$ $\left(\left[\begin{smallmatrix}1&0&1\\0&0&0\\0&0&0\end{smallmatrix}\right], \left[\begin{smallmatrix}1&0&1\\0&0&0\\0&0&0\end{smallmatrix}\right], \left[\begin{smallmatrix}1&0&1\\0&0&0\\0&0&0\end{smallmatrix}\right] \right),$ $\left(\left[\begin{smallmatrix}1&1&0\\1&0&0\\0&1&0\end{smallmatrix}\right], \left[\begin{smallmatrix}1&1&0\\1&0&0\\0&1&0\end{smallmatrix}\right], \left[\begin{smallmatrix}1&1&0\\1&0&0\\0&1&0\end{smallmatrix}\right] \right),$ $\left. \left(\left[\begin{smallmatrix}0&0&0\\0&1&0\\0&1&0\end{smallmatrix}\right], \left[\begin{smallmatrix}0&0&0\\0&1&0\\0&1&0\end{smallmatrix}\right], \left[\begin{smallmatrix}0&0&0\\0&1&0\\0&1&0\end{smallmatrix}\right] \right) \right\}$
$\left\langle \triangle, \phi_{\left[\begin{smallmatrix}0&1&0\\1&0&0\\0&0&1\end{smallmatrix}\right],I,I} \circ \top \right\rangle$	not conjugate to G_{Ballard}	$\left\{ \left(\left[\begin{smallmatrix}0&1&1\\1&1&1\\1&1&1\end{smallmatrix}\right], \left[\begin{smallmatrix}1&1&1\\0&1&1\\1&1&1\end{smallmatrix}\right], \left[\begin{smallmatrix}1&1&1\\1&0&1\\1&1&1\end{smallmatrix}\right] \right), \right.$ $\left(\left[\begin{smallmatrix}0&0&0\\1&0&1\\1&0&1\end{smallmatrix}\right], \left[\begin{smallmatrix}0&0&0\\0&1&1\\0&1&1\end{smallmatrix}\right], \left[\begin{smallmatrix}1&0&1\\0&0&0\\1&0&1\end{smallmatrix}\right] \right),$ $\left(\left[\begin{smallmatrix}0&0&0\\1&1&1\\1&1&1\end{smallmatrix}\right], \left[\begin{smallmatrix}1&0&1\\0&1&1\\1&1&1\end{smallmatrix}\right], \left[\begin{smallmatrix}1&0&1\\1&0&1\\1&0&1\end{smallmatrix}\right] \right),$ $\left(\left[\begin{smallmatrix}1&1&1\\1&1&1\\1&1&1\end{smallmatrix}\right], \left[\begin{smallmatrix}1&1&1\\1&1&1\\1&1&1\end{smallmatrix}\right], \left[\begin{smallmatrix}1&1&1\\1&1&1\\1&1&1\end{smallmatrix}\right] \right),$ $\left. \left(\left[\begin{smallmatrix}1&0&1\\0&0&0\\0&0&0\end{smallmatrix}\right], \left[\begin{smallmatrix}1&0&1\\0&0&0\\0&0&0\end{smallmatrix}\right], \left[\begin{smallmatrix}1&0&1\\0&0&0\\0&0&0\end{smallmatrix}\right] \right) \right\}$
		$\left\{ \left(\left[\begin{smallmatrix}1&0&1\\1&1&1\\1&1&1\end{smallmatrix}\right], \left[\begin{smallmatrix}0&1&1\\1&1&1\\1&1&1\end{smallmatrix}\right], \left[\begin{smallmatrix}1&1&1\\1&0&1\\1&1&1\end{smallmatrix}\right] \right), \right.$ $\left. \left(\left[\begin{smallmatrix}0&0&0\\0&1&1\\0&1&1\end{smallmatrix}\right], \left[\begin{smallmatrix}0&1&1\\0&0&0\\0&1&1\end{smallmatrix}\right], \left[\begin{smallmatrix}1&0&1\\0&0&0\\1&0&1\end{smallmatrix}\right] \right), \right.$

		$\left(\begin{bmatrix}0&0&0\\1&1&1\\1&1&1\end{bmatrix},\begin{bmatrix}0&1&1\\1&0&1\\1&1&1\end{bmatrix},\begin{bmatrix}1&0&1\\1&0&1\\1&0&1\end{bmatrix}\right),$ $\left(\begin{bmatrix}1&1&1\\1&1&1\\1&1&1\end{bmatrix},\begin{bmatrix}1&1&1\\1&1&1\\1&1&1\end{bmatrix},\begin{bmatrix}1&1&1\\1&1&1\\1&1&1\end{bmatrix}\right),$ $\left(\begin{bmatrix}0&1&1\\0&0&0\\0&0&0\end{bmatrix},\begin{bmatrix}0&0&0\\1&0&1\\0&0&0\end{bmatrix},\begin{bmatrix}1&0&1\\0&0&0\\0&0&0\end{bmatrix}\right)\}$

$S = \{\triangle, \top\}$:

$\langle F\rangle = \ldots$	conjugacy	$D = \langle F\rangle(E)$, where $E = \ldots$
$\left\langle \triangle, \top, \phi_{\begin{bmatrix}0&1&0\\1&0&0\\0&0&1\end{bmatrix},\begin{bmatrix}0&1&0\\1&0&0\\0&0&1\end{bmatrix},I} \right\rangle$	$= \left(\phi_{\begin{bmatrix}1&0&1\\0&1&1\\1&1&1\end{bmatrix},\begin{bmatrix}1&0&1\\0&1&1\\1&1&1\end{bmatrix},\begin{bmatrix}1&0&1\\0&1&1\\1&1&1\end{bmatrix}}\right) \star \left\langle \triangle, \phi_{\begin{bmatrix}0&1&0\\1&0&0\\0&0&1\end{bmatrix},I,I} \circ \top \right\rangle$	$\left\{\left(\begin{bmatrix}0&0&0\\0&0&0\\0&0&1\end{bmatrix},\begin{bmatrix}0&0&0\\0&0&0\\0&0&1\end{bmatrix},\begin{bmatrix}0&0&0\\0&0&0\\0&0&1\end{bmatrix}\right),\right.$ $\left(\begin{bmatrix}1&0&0\\0&0&0\\0&0&0\end{bmatrix},\begin{bmatrix}1&0&0\\0&0&0\\0&0&0\end{bmatrix},\begin{bmatrix}1&0&0\\0&0&0\\0&0&0\end{bmatrix}\right),$ $\left(\begin{bmatrix}0&1&1\\0&0&0\\0&0&0\end{bmatrix},\begin{bmatrix}0&0&0\\1&0&1\\0&0&0\end{bmatrix},\begin{bmatrix}1&0&1\\0&0&0\\0&0&0\end{bmatrix}\right),$ $\left(\begin{bmatrix}0&1&1\\0&0&0\\0&1&0\end{bmatrix},\begin{bmatrix}0&0&0\\1&0&1\\1&0&0\end{bmatrix},\begin{bmatrix}1&0&1\\0&0&0\\1&0&0\end{bmatrix}\right),$ $\left.\left(\begin{bmatrix}0&1&1\\1&0&1\\1&1&1\end{bmatrix},\begin{bmatrix}0&0&0\\0&0&0\\1&0&0\end{bmatrix},\begin{bmatrix}0&0&1\\0&0&0\\0&0&0\end{bmatrix}\right)\right\}$
		$\left\{\left(\begin{bmatrix}0&0&0\\0&0&0\\0&0&1\end{bmatrix},\begin{bmatrix}0&0&0\\0&0&0\\0&0&1\end{bmatrix},\begin{bmatrix}0&0&0\\0&0&0\\0&0&1\end{bmatrix}\right),\right.$ $\left(\begin{bmatrix}0&1&0\\0&0&0\\0&0&0\end{bmatrix},\begin{bmatrix}0&0&0\\1&0&0\\0&0&0\end{bmatrix},\begin{bmatrix}1&0&0\\0&0&0\\0&0&0\end{bmatrix}\right),$ $\left(\begin{bmatrix}1&0&1\\0&0&0\\0&0&0\end{bmatrix},\begin{bmatrix}1&0&1\\0&0&0\\0&0&0\end{bmatrix},\begin{bmatrix}1&0&1\\0&0&0\\0&0&0\end{bmatrix}\right),$ $\left(\begin{bmatrix}1&0&1\\0&0&0\\1&0&0\end{bmatrix},\begin{bmatrix}1&0&1\\0&0&0\\1&0&0\end{bmatrix},\begin{bmatrix}1&0&1\\0&0&0\\1&0&0\end{bmatrix}\right),$ $\left.\left(\begin{bmatrix}1&0&1\\0&1&1\\1&1&1\end{bmatrix},\begin{bmatrix}0&0&0\\0&0&0\\1&0&0\end{bmatrix},\begin{bmatrix}0&0&1\\0&0&0\\0&0&0\end{bmatrix}\right)\right\}$
$\left\langle \triangle, \top, \phi_{\begin{bmatrix}0&1&0\\1&1&0\\0&0&1\end{bmatrix},\begin{bmatrix}0&1&0\\1&1&0\\0&0&1\end{bmatrix},\begin{bmatrix}0&1&0\\1&1&0\\0&0&1\end{bmatrix}} \right\rangle$	not conjugate to any other subgroup shown here; solution has rank 27	$\left\{\left(\begin{bmatrix}0&0&0\\1&0&1\\1&0&1\end{bmatrix},\begin{bmatrix}0&0&0\\1&0&1\\1&0&1\end{bmatrix},\begin{bmatrix}0&0&0\\1&0&1\\1&0&1\end{bmatrix}\right),\right.$ $\left(\begin{bmatrix}0&1&0\\1&0&1\\1&0&1\end{bmatrix},\begin{bmatrix}1&0&1\\1&0&1\\1&0&1\end{bmatrix},\begin{bmatrix}1&0&1\\0&0&0\\0&0&0\end{bmatrix}\right),$ $\left(\begin{bmatrix}0&0&0\\1&1&1\\1&1&1\end{bmatrix},\begin{bmatrix}0&1&1\\0&1&1\\0&1&1\end{bmatrix},\begin{bmatrix}1&0&1\\0&0&0\\1&0&1\end{bmatrix}\right),$ $\left.\left(\begin{bmatrix}0&0&0\\0&1&1\\0&1&1\end{bmatrix},\begin{bmatrix}1&1&1\\0&0&0\\1&1&1\end{bmatrix},\begin{bmatrix}1&0&1\\1&0&1\\1&0&1\end{bmatrix}\right)\right\}$

Our source code is available at

`https://github.com/coolcomputery/Matrix-Multiplication-Tensor-Decomposition`.

5.1 Using flip graphs

[Kauers & Moosbauer, 2022] recently published their "flip graph" method to find low-rank mod 2 decompositions of $\mathcal{T}^{\langle n,k,m\rangle}$, which they used to obtain $\langle 4,4,4:47\rangle_{\mathbb{Z}_2}$ (matching AlphaTensor [Fawzi et al., 2022]) and $\langle 5,5,5:95\rangle_{\mathbb{Z}_2}$. The method repeatedly applies 'flips' (which change two matrix triplets at a time) and 'reductions' (which replace a subset of n triplets with $n-1$ triplets if a certain linear dependence condition is met) in a randomized manner, effectively taking a random walk in the graph where decompositions are nodes and flip/reductions are edges (the 'flip graph'). [10]

We ran the provided source code of the flip graph method on all the decompositions of $\mathcal{T}^{\langle 3\rangle}$ that we found, specifying argument values pathlength=10^9, restart=1. The decomposition with rank 27 was only reduced to rank 23 and took about 6 hours of computation. For every other decomposition, the flip graph algorithm terminated immediately without produc-

[10] The flip graph does in fact account for symmetries of $\mathcal{T}^{\langle n\rangle}$: it uses the equivalence class where any two decompositions that can be transformed to each other by applying some $f \in \mathbf{\Gamma}$ are equivalent. [Kauers & Moosbauer, 2022]

ing any output; adding some console-printing lines in the code revealed that the algorithm was unable to do a flip or a reduction on the first step of the walk, suggesting that every such decomposition is an isolated point in the flip graph.

To confirm that this is indeed the case, we note that a flip operation transforms $\{(A,B,C),(A,B',C')\}$ to $\{(A,B+B',C),(A,B',C'-C)$ (or any index-permuted variant), so it requires both triplets to share a matrix at some index; and a reduction operation on triplets $\left\{\left(A^{(k)},B^{(k)},C^{(k)}\right)\right\}_k$ requires that $\left\{A^{(k)}\right\}_k$, $\left\{B^{(k)}\right\}_k$, or $\left\{C^{(k)}\right\}_k$ span a 1-dimensional subspace, i.e. for one of those sets to consist of matrices that are just different scalar multiples of the same underlying matrix. For each decomposition D we found that had rank 23, and for every index $0 \le a < 3$, the matrices d_a for all $d \in D$ (i.e. the a^{th} matrix of every triplet in D) were all nonzero and mutually distinct; this immediately means no flips are possible, and since we are working over $\mathbb{Z}_2$, no reductions are possible either. For the last decomposition we found that had rank 27, $|\{d_a : d \in D\}| = 21$ for each a, which explains why it could be improved with flips and reductions.

5.2 Takeaways from search results

We still have not found a decomposition of $\mathcal{T}^{\langle 3\rangle}$ with rank 22 or lower. However, we note some interesting observations:

- Among all rank ≤ 23 decompositions of $\mathcal{T}^{\langle 3\rangle}$ we found, there are four distinct symmetry subgroups, but only two distinct subgroups up to conjugacy, with one of them being conjugate to G_{Ballard}.
- Since none of the decompositions we found have rank < 23, by Corollary 1 this means that no such decompositions exist (over $\mathbb{Z}_2$) for any of the symmetry subgroups listed above, so analyzing them further would be a dead end.
- Setting $F = \{\text{id}\}, R = 2$ yields no solutions, so the rank of $\mathcal{T}^{\langle 3\rangle}$ is at least 3 over $\mathbb{Z}_2$. This is a very weak lower bound; in fact, the rank of $\mathcal{T}^{\langle 3\rangle}$ is known to be at least 19 over arbitrary fields. [Bläser, 2003]
- For $S = \{\triangle, \top\}$, every symmetry subgroup $\langle S \cup \{f\}\rangle$ except for $\langle \triangle, \top, \text{id}\rangle$ has been proven to not have a rank < 23 decomposition.
- For $S = \{\triangle\}$, if we restrict our attention to rank ≤ 21 decompositions (since 21 is the largest rank for which we would surpass Strassen's algorithm, as $\log_3 21 \approx 2.771 < 2.807 \approx \log_2 7$), the only subgroups $\langle S \cup \{f\}\rangle$ that might have a rank ≤ 21 decomposition are $\langle \triangle, \text{id}\rangle$, $\langle \triangle, \top\rangle$, and $\left\langle \triangle, \phi_{M=\left[\begin{smallmatrix}1&1&0\\0&1&0\\0&0&1\end{smallmatrix}\right],M,M} \right\rangle$. The last of these subgroups seems to be the most feasible to analyze and might be worth investigating further.

6 Future directions

Despite our unsuccessful attempt to obtain a rank-22 decomposition of $\mathcal{T}^{\langle 3\rangle}$ with our algorithm and with flip graphs, we still think that combining these two approaches is promising and that there is still much left to uncover; although [Kauers & Moosbauer, 2022] account for symmetries of $\mathcal{T}^{\langle n\rangle}$ in their equivalence classes between decompositions, they do not seem to investigate individual *decompositions* that have high amounts of symmetry.

Another strategy worth attempting could be generalizing flips and reductions to more complex transformations of decompositions, since all of our solutions except the last one with rank 27 could not be transformed at all with flips and reductions alone.

Finally, bigger tensors such as $\mathcal{T}^{\langle 4\rangle}$ can be investigated for symmetric decompositions, although extra restrictions may have to be enforced on what matrix triplets are allowed, since enumerating all matrix triplets ($\approx 2^{(3n^2)}$ of them) is only feasible on a home computer up to $n = 3$. One such restriction could be a sparsity constraint, such as setting a maximum allowed value of $\mathtt{b}(A)\,\mathtt{b}(B)\,\mathtt{b}(C)$ for each triplet (A,B,C), where $\mathtt{b}(M)$ is the number of 1s in matrix M.

7 Acknowledgments

We thank Prof. Virginia Vassilevska Williams for her support throughout the past year-and-a-half, and for her suggestions to work over $\mathbb{Z}_2$ and to test symmetry groups that superset $\{\triangle, \top\}$.

References

Strassen, V. (1969). Gaussian elimination is not optimal. *Numerische Mathematik*, 13, pp. 354–356. Retrieved on March 24, 2023 from https://doi.org/10.1007/BF02165411

Landsberg, J. M. (2019). Efficient matrix multiplication. *ILAS July 2019*. Retrieved on March 30, 2023 from https://www.math.tamu.edu/~jml/ILAS7-19.pdf

Bläser, M. (2013). Fast Matrix Multiplication. *Theory of Computing Library*, Graduate Surveys 5, pp. 1-60. Retrieved on March 24, 2023 from https://theoryofcomputing.org/articles/gs005/

De Groote, H. F. (1978). On varieties of optimal algorithms for the computation of bilinear mappings I. the isotropy group of a bilinear mapping. Retrieved on March 30, 2023 from https://doi.org/10.1016/0304-3975(78)90038-5

Ballard, G., Ikenmeyer, C., Landsberg, J. M., & Ryder, N. (2018). The Geometry of Rank Decompositions of Matrix Multiplication II: 3 × 3 Matrices. Retrieved on March 24, 2023 from https://arxiv.org/abs/1801.00843

Duan, R., Wu, H., & Zhou, R. (2022). Faster Matrix Multiplication via Asymmetric Hashing. Retrieved on March 25, 2023 from https://arxiv.org/abs/2210.10173

Smirnov, A. V. (2013). The Bilinear Complexity and Practical Algorithms for Matrix Multiplication. *ISSN 09655425, Computational Mathematics and Mathematical Physics*, 2013, Vol. 53, No. 12, pp. 1781–1795. Retrieved on March 24, 2023 from https://cs.uwaterloo.ca/~eschost/Exam/Smirnov.pdf

Laderman, J. D. (1976). A Non-Commutative Algorithm for Multiplying 3x3 Matrices Using 23 Multiplications. *Bulletin of the American Math Society*. 82(1), 126-128. Retrieved on March 24, 2023 from https://www.ams.org/journals/bull/1976-82-01/S0002-9904-1976-13988-2/S0002-9904-1976-13988-2.pdf

Pan, V. (1978). Strassen's algorithm is not Optimal Trilinear Technique of Aggregating, Uniting and Canceling for Constructing Fast Algorithms for Matrix Operations. *Proceedings of the 19th Annual Symposium on Foundations of Computer Science*. Retrieved on March 25, 2023 from https://dl.acm.org/doi/10.1109/SFCS.1978.34

Pan, V. (1982). Trilinear aggregating with implicit canceling for a new acceleration of matrix multiplication. Retrieved on March 30, 2023 from https://doi.org/10.1016/0898-1221(82)90037-2

Kolda, T. & Bader, B. W. (2009). Tensor Decompositions and Applications. Retrived on March 25, 2023 from https://doi.org/10.1137/07070111X.

Bläser, M. (2003). On the complexity of the multiplication of matrices of small formats. *Journal of Complexity*, 19, 43-60. Retrieved on March 24, 2023 from https://doi.org/10.1016/S0885-064X(02)00007-9.

Fawzi, A., Balog, M., Huang, A., Hubert, T., Romera, B., Barekatain, M., Novikov, A., Ruiz, F. J. R., Schrittwieser, J., Swirszcz, G., Silver, D., Hassabis, D., & Kohli, P., (2022). Discovering faster matrix multiplication algorithms with reinforcement learning. Retrieved on March 25, 2023 from https://www.nature.com/articles/s41586-022-05172-4

Kauers, M. & Moosbauer, J. (2022). Flip Graphs for Matrix Multiplication. Retrieved on March 25, 2023 from https://arxiv.org/abs/2212.01175

At Pfizer, we pursue groundbreaking science every single day, fueled by one purpose: to develop **breakthroughs that change patients' lives**. We deliver on this purpose by building courageous, diverse teams of experts who contribute to all aspects of what we do and providing them with state-of-the-art resources, technology, and facilities.

Discover how you can help ensure that Science Will Win – including opportunities in Cambridge – at **careers.pfizer.com**.

Pfizer


YOUR FUTURE STARTS WITH US
HRL
LABORATORIES
FOR MORE THAN 70 YEARS, HRL'S SCIENTISTS AND ENGINEERS HAVE BEEN ON THE LEADING EDGE OF TECHNOLOGY, CONDUCTING PIONEERING RESEARCH, PROVIDING REAL-WORLD TECHNOLOGY SOLUTIONS, AND ADVANCING THE STATE OF THE ART.
With more than 100 positions available
Find your future at www.hrl.com/careers