BAR-ILAN UNIVERSITY–YESHIVA UNIVERSITY

Summer Science Research Internship Program 2024



The Bar-Ilan University–Yeshiva University Summer Science Research Internship Program is an amazing research opportunity for undergraduate men and women, allowing them to contribute to the forefront of science research taking place in Israel. Generously supported by former chairman of Bar-Ilan's Global Board of Trustees, Dr. Mordecai D. Katz z"l and his wife Dr. Monique Katz, the Irving I. Stone Foundation and the Zoltan Erenyi Fund, students gain invaluable laboratory skills, along with an unforgettable summer experience.

Program Directors: Prof. Arlene Wilson-Gordon and Prof. Ari Zivotofsky
Av and Em Bayit: Rabbi Feivel and Tamar Segelov

TABLE OF CONTENTS

Brain Sciences.....4

Can Social Anxiety Disorder be Detected in Linguistic Features of Autobiographical Narratives?.....4
Mynda Barenholtz and Emuna Gozland

Language Coactivation: Using Eye Tracking Technology to Explore Coactivation Effects in Trilingual Individuals.....5
Aryeh Gross and Eliana Diamond

The Effect of Gratitude Journaling on the Anticipation-Memory-Experience Gap.....8
Ruben Prawer and Isaac Dietz

Comparative Gene Expression in Nocturnal Aplysia fasciata and Diurnal Aplysia californica: C/EBP and Creb1a in Memory Consolidation.....9
Sarah Grunseid

Analyzing Alzheimer's Disease: Measuring Amyloid-Beta in the Brains of Transgenic 5xFAD Mouse Models.....11
Esther Nahon

Engineering.....13

BiCCA Adaption of Canonical Correlation Analysis.....13
Jordan Paraboschi

Connecting and Evaluating the Reasoning Engine on a Repository of real-world Boolean Network Models.....15
Blimy Dresdner and Hannah Zahler

Measuring Blood Pressure Using Light.....17
Betzalel Cohen

Computational Studies of PbSe Metasurface for Mid-Infrared Metaphotonics.....19
Yishai Kohn

Multi-Agent Constraint CBS for Multi-Agent Path Finding.....22
Tuvya Macklin

Exploring Single Photon Emitters in Hexagonal Boron Nitride for Quantum Applications.....24
David Benhamu

Detection of Atto 532 Fluorescent Dye Using a High Throughput Optical Modulation Biosensing System.....26
Perie Hoffman

Retention Time Tests on Gain Cell Embedded Dynamic Random Access Memory.....27
Hannah Bialik

Life Sciences.....30

Genome Editing by CRISPR-Cas9 System for Modeling and Correction of Hematological Diseases.....30
Tobi Katz

HSV-1 Infection Expression on Cellular Factors.....32
Hannah Mamet

Survival Outcomes and Prognostic Factors in Acute Myeloid Leukemia Patients Through Rigorous Data Analysis.....34
Lilly Glickman

Determining the Role of Codanin-1 in Congenital Dyserythropoietic Anemia Type I.....37
Meira Strauss

A Study of Sirt6, the Anti-aging Protein, and its Overexpression in Liver Tissue.....39
Ella Nasser and Alyssa Polonetsky

Physics, Chemistry, and Mathematics.....42

The Varieties of n-Dimensional Lie Algebras and Their Components.....42
Jonathan Beer

Discrete Curve Shortening Flows and Geometric Properties through Fourier Decomposition of Planar Polygons.....43
Daniel Lipshitz

Machine Learning-Guided Discovery of Multinary Metal Oxides for Sustainable Energy Applications. 45
Rachael Farkas

Inhibition of P53 aggregation with cyclic D,L-α-peptide.....46
Eliana Fishman

Platinum Nanoparticles Catalyst for Oxygen Reduction Reaction in Hydrogen Fuel Cells.....48
Elana Rosenblatt

Acoustic Analogues of Black Holes and Gravitational Waves.....50
Sophia Purow

Editor: Jonathan Beer
Contributing Editor: Eliana Diamond

BRAIN SCIENCES



Ruben Prawer, Isaac Dietz, Aryeh Gross,
Esther Nahon, Emuna Gozland, Sarah Grunseid, Mynda Barenholtz, Eliana Diamond

Can Social Anxiety Disorder be Detected in Linguistic Features of Autobiographical Narratives?

Mynda Barenholtz and Emuna Gozland
Advised under Prof. Eva Gilboa-Schechtman

Social Anxiety Disorder is a mental health condition marked by an intense fear of social situations and being judged by others, leading to significant distress and avoidance of social interactions. This condition can interfere with daily activities and relationships[1]. The Liebowitz Social Anxiety Scale (LSAS) is a well-established tool that measures social anxiety by assessing fear and avoidance across various social situations and interactions. While the LSAS effectively quantifies social anxiety in specific situations, it is unclear if these features appear in the way individuals narrate their experiences of social exclusion, humiliation, and pride.

Research indicates that individuals with social anxiety disorder tend to have negative self-beliefs and a distorted self-view when recalling

painful autobiographical social memories. This is reflected in their language use, as they tend to use more self-referential and anxiety-related words and make fewer references to others[2]. Additionally, sentiment analysis has been employed to detect PTSD from text data, showing the potential for machine learning models to identify emotional distress from narrative content[3]. However, there is a gap in research regarding how the narrative expression of social anxiety features correlates with standardized measures like the LSAS.

Our methods include coding previously collected autobiographical stories from participants (n≈1700) and coding them for narrative features such as specificity, detail, emotional intensity, and self-evaluation. Multiple coders used a detailed coding key to evaluate the narratives, while an AI model independently coded the same narratives. We will then compare the AI's coding with the human coding to assess the accuracy and consistency of the AI model. The AI will iteratively check its reliability against the human coders and adjust its algorithms based on this feedback until it reaches a stable and accurate level of reliability, and can code on its own.

This study could provide a novel method for detecting social anxiety through narrative analysis of how individuals linguistically code and remember autobiographical memories of humiliation, exclusion, and pride. We anticipate finding a significant correlation between the narrative features and the LSAS scores, suggesting that individuals with higher social anxiety scores will exhibit distinct patterns in their narratives. This approach has the potential to aid in the diagnosis and understanding of social anxiety by revealing how it influences the recollection and communication of social experiences.

[1] American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.).

[2] Anderson, B., Goldin, P. R., Kurita, K., & Gross, J. J. (2008). Self-representation in social anxiety disorder: Linguistic analysis of autobiographical narratives. *Behaviour Research and Therapy, 46*(10), 1119-1125.

[3] Sawalha, J., Yousefnezhad, M., Shah, Z., Brown, M. R. G., Greenshaw, A. J., & Greiner, R. (2022). Detecting presence of PTSD using sentiment analysis from text data. *Journal of Affective Disorders, 309,* 287-295.

Language Coactivation: Using Eye Tracking Technology to Explore Coactivation Effects in Trilingual Individuals

Aryeh Gross and Eliana Diamond
Advised under Professor Natalia Meir and Professor Sharon Armon-Lotem

Israel is unique as it contains a large population of Russian immigrants with proficiency in English, Hebrew, and Russian. The study by Fridman and Meir[1] uses eye tracking technology to track the differences in eye-fixation as well as reaction times between trilingual English-Hebrew-Russian speakers compared to bilingual and monolingual English, Hebrew, and Russian speakers. The data for monolingual English speakers were collected in the current project.

As our world becomes even more connected and immigration becomes more ubiquitous, studies such as these are important to better understand the effects of knowing multiple languages and its impact on how people interact with the world around them.

While previous research has shown language coactivation in bilingual individuals[2], the study by Fridman and Meir[1] aimed to investigate phonological cross-linguistic coactivation during word recognition in trilingual speakers of Russian (Heritage Language), Hebrew (Societal Language), and English (L3) in their L3

compared to monolingual L1-English speakers. In the present project, we collected data for monolingual English-speaking controls. The participants in this study consisted of trilingual English-Hebrew-Russian speakers[1] and a control group of monolingual English (the current project).

Participants completed a Multilingual Naming Test (MINT) individually before completing the object-identification task. In this task, participants were shown a series of images and were asked to identify the objects as they appeared. This test established the baseline for participants' ability to recognize and label objects for the second part of the experiment. The MINT assessment[3] was explicitly designed to test multilinguals in contrast to previous picture-naming vocabulary assessments such as the Boston Naming Test, which had been developed with only monolinguals in mind. The MINT contains 67 black and white line drawings, prompting 68 picture names (one image prompts two lexical labels: mortar and pestle) in increasing difficulty.

Table 1. MINT Test Scores in Trilingual v Monolingual Speaker

Data from Fridman and Meir (under review)	**Data collection in the current project**
Trilingual L1-Russian/ L2-Hebrew/ L3-English speakers	Monolingual L1-English speakers
n=48	n=9
MINT Scores in L3-English M=76% (SD=9%)	MINT Scores for Monolingual L1-English M=91.40% (SD=5%)
Data from Fridman and Meir (under review)	Data collection in the current project

After completing the MINT test, participants were seated by an EyeLink eye tracking device and had the experimental task explained to them. The participants were shown a series of 42 slides each containing four images of objects while the computer prompted them to select a specific object.

Within the English condition, where the instructions were presented in English, there were four conditions of slides presented to the participants. The first condition contained images with no overlapping phonological features (Table 2A) and would therefore not elicit any language coactivation. The second and third conditions showed images with coactivation in Hebrew (Table 2B) and Russian (Table 2C) when the experiment was conducted using English audio instructions. The last condition presented a three language coactivation for both L1-Russian, L2-Hebrew, and L3-English, with phonologically similar sounds across the three languages (Table 2D).

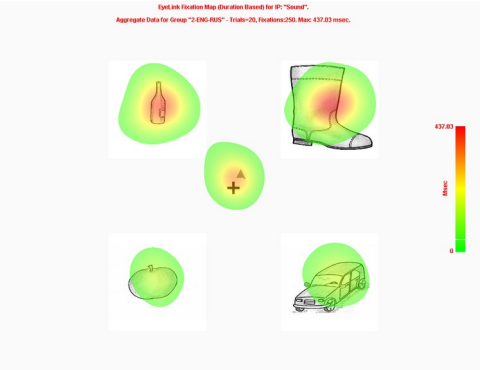
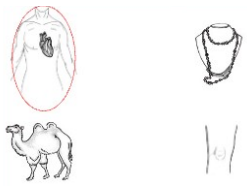
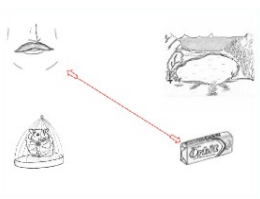
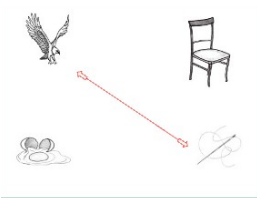
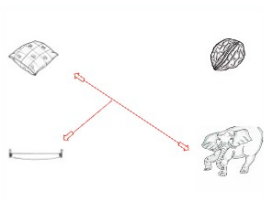


Figure 1: EyeLink Fixation Map (2 language coactivation ENG-RUS: butulka 'bottle.Russian' and boot)

Table 2. Conditions of the study

(A)	(B)	(C)	(D)
No Activation (L1-English)	Two Language Coactivation (Hebrew and English)	Two Language Coactivation (Russian and English)	Three Language Activation (L1-Russian, L2-Hebrew, L3-English)
Where is the heart?	Where is the **mus**tache? ***mas**tik* (gum.Hebrew)	Where is the **eagl**e? ***igl**a* (needle.Russian)	Where is the **pill**ow? ***pil**a* (saw.Russian) ***pil*** (elephant.Hebrew)
			

Through this, we can analyze the differences in eye-movement between the different groups to see the effects of the cross-language phonetic similarity for trilinguals compared to bilinguals and monolinguals. The results from Fridman and Meir[1] are presented in Figure 2 for the English data.

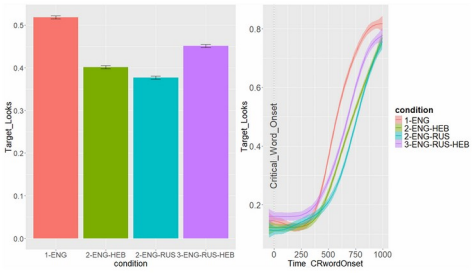


Figure 2: English Experiment Mean Eye-Fixation Proportions and Time Course Analyses

Within the Russian experiment, both Hebrew and English-Hebrew competitors showed significant delays in reaction times compared to the Russian-only condition. Within the Hebrew experiment, almost no significant slow-downs in reaction times were observed. This delay in reaction time is corroborated by how the participant's eye movements were shown to move back and forth between objects. In contrast, slow-downs occurred in every condition within the English experiment. The strongest reaction-time delay occurred in the English-Hebrew condition, but there were still effects in the English-Russian, and English-Russian-Hebrew distractor conditions[1].

The delays in reaction times indicate that the other languages are always "running in the background" for those who know multiple languages. While the time delays in selecting the images for trilinguals imply a negative consequence of knowing multiple languages, as it may require more effort and time to communicate during the day-to-day, this phenomenon of the other languages "running in the background" also provides trilinguals with an advantage. Trilingual individuals continually exercise all their proficient languages, even when communicating in only one of them. While interacting in one language, the other

two run continuously, therefore challenging and exercising the brain. In this case, trilingual individuals keep their brain active and healthy just by living their lives.

[1] Fridman, C. & Meir, N. (under review)

[2] W. Szubko-Sitarek, Studies in second language Learning and Teaching, 21, 189-208 [2011]

[3] T.H. Gollan et al., Bilingualism: Language and Cognition, 159(3):594-615 [2012]

The Effect of Gratitude Journaling on the Anticipation-Memory-Experience Gap

Ruben Prawer and Isaac Dietz
Advised under Prof. Shimrit Daches

One's emotions toward an experience can be broken down into three parts: anticipation, experience, and memory. Anticipation refers to one's predictions of their future emotions when encountering the event; experience is one's emotional state during the actual event; and memory refers to one's recalled emotions when looking retrospectively. These different parts of experience influence each other. For example, recalled emotions have a significant impact on how one predicts their emotions in a similar future situation. However, there are distinctions between the components of experience. An important example is that we typically predict and recall both positive and negative emotions as more intense than we experienced them in the moment[1]. This discrepancy between predicted/recalled emotions and experienced emotions is referred to as the Anticipation-Experience-Memory gap.

Furthermore, the connections and distinctions between the parts of experience are moderated by inter-individual differences. These individual differences are observed with people with depression in a concept called depressive realism, which refers to how depressed individuals are often more accurate in their prediction of their future emotional state than people without depression. Other individual differences are seen with individuals with schizophrenia and anhedonia.

Gratitude involves noticing, appreciating, and savoring the positive elements of one's life. It is correlated with greater positive emotions, counteracts negative emotions, and is related to positive social outcomes. Gratitude journaling (GJ) is a proven method of increasing trait gratitude that involves participants writing down events, happenings, or people to whom they are grateful on a regular basis. GJ has been shown to increase positive emotionality, decrease negative emotionality, and have positive physiological effects such as reducing stress and the risk of heart failure[2][3]. Furthermore, GJ creates a positive feedback loop: GJ causes more positive attitudes, which leads to positive feelings, and these feelings result in adaptive and beneficial behavior that culminates in more positive feelings.

Our lab is interested in manipulating gratitude through GJ and observing if this manipulation affects the Anticipation-Experience-Memory gap. To test this, we plan first to run a pilot study that does not manipulate gratitude but instead measures natural variation in gratitude among the population and observe if trait gratitude is correlated with changes in the Anticipation-Experience-Memory gap. The method would involve using a questionnaire to test trait gratitude and then measuring predicted, experienced, and recalled emotions for a specific experience, in this case, a weekend. Predicted and recalled emotions will be measured using a typical survey; however, we will use an ecological momentary assessment to measure experienced emotion. Assuming that we find that trait gratitude has an effect, we will move on to our main study. The study will manipulate trait gratitude by

having one group perform GJ for three weeks while the neutral group journals about non-gratitude-related items. Additionally, there will be a control group that does not journal. The measures of predicted, experienced, and remembered emotions will follow the same procedure as the pilot study.

We predict one of two possible outcomes. The first possibility is that GJ will cause an increase in intensity of recalled emotions as journaling affects how one looks back on experiences. This is because GJ engenders a positive retrospective outlook. In turn, the increased recalled emotionality will cause an increased intensity of predicted emotion. However, the intensity of experienced emotion may be unchanged; thus, we will find that the Anticipation-Experience-Memory gap will increase amongst GJ participants. Alternatively, GJ may affect how one actually experiences the event and, therefore, increase the intensity of emotionality during the experience. In this hypothesis, the Anticipation-Experience-Memory gap will remain the same for both groups as the GJ group displays increased emotionality for all three components of the experience.

The implications of this research are twofold. Firstly, it adds to the growing literature on GJ. Additionally, these results may bear important findings for people with depression or other emotional disorders that may benefit from GJ interventions to assist them in developing a more adaptive Anticipation-Experience-Memory gap.

[1] D. Wirtz et al., Psychological Science, 14, 520-524, (2003).

[2] S. Işık and B. Ergüner-Tekinalp, International Journal for the Advancement of Counselling, 39, 164–175, (2017)

[3] L. Redwine et al., Psychosom Med., 78, 667-676, (2016).

Comparative Gene Expression in Nocturnal Aplysia fasciata and Diurnal Aplysia californica: C/EBP and Creb1a in Memory Consolidation

Sarah Grunseid
Advised under Prof. Avy Susswein, Dr. Yisrael Schnytzer, and Master's Student Yarden Marom

The sea slugs, *Aplysia*, have long served as a model organism for studying learning and memory due to the simplicity of their nervous system and the presence of large, identifiable neurons. Nobel prize winner Eric Kandel's extensive research on *Aplysia* has shed light on the differences between short term and long term memory. Unlike short term memory, long term memory modifies gene expression, thereby stimulating protein synthesis and generating new synapses. Additionally, Kandel delineated the memory consolidation process: in which sensory input triggers cerebral transcription followed by hippocampus activation and ultimately leads to storage of memories in the cortex[1]. The initial steps occur during waking hours while the final steps are paused and completed during sleep. Previous studies suggest that when *Aplysia* are presented with new stimuli during their putative resting period, a natural memory blocker is activated to prevent the new information from interfering with the prior ongoing memory consolidation process. Furthermore, the genes, *C/EBP* and *Creb1a* are believed to play significant roles in these processes of memory formation and consolidation[2]. Our study tries to elucidate the expression level of these genes when diurnal and nocturnal slugs' sleep is disturbed; and attempts to further clarify whether levels of expression of *C/EBP* and *Creb1a* are actually

correlated to wake and sleep cycle, memory consolidation and or lip stimulation.

We carried out various experiments utilizing the inedible food task in order to promote learning and memory consolidation in the *Aplysia*. The task involved feeding the animals *Ulva* enclosed in a net, leading to repeated but unsuccessful attempts at swallowing. Over time, the animals made fewer attempts at ingestion, indicating successful memory formation. Past studies have established that a 15-minute training session is adequate for memory formation. Thus the experimental group was trained using the inedible food task for 15-minutes. To control for potential gene expression changes caused solely by lip stimulation, a 15-minute lip stimulation was performed on a subset of animals. Additionally, a third group of naive animals were used as a control. Resting periods of 15 and 120 minutes were incorporated to account for immediate and delayed gene expression changes, respectively. Treatments were performed on the diurnal and nocturnal animals trained both during the day and during their putative resting period. Following the training sessions, the animals were sacrificed and their Buccal and Cerebral ganglia were removed for total RNA extraction and cDNA synthesis. Quantitative PCR was then performed on both ganglia for all treatments on the genes Creb1a and C/EBP, using GAPDH as a housekeeping gene.

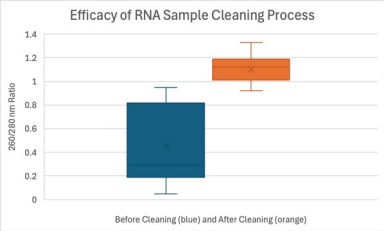
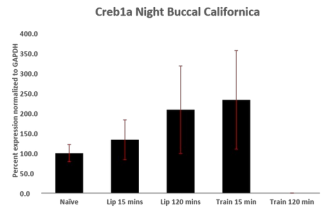
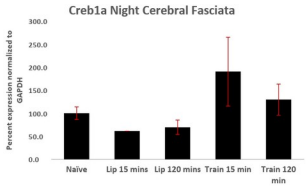


Figure 1: Comparison of RNA Purity Before and After Cleaning Process

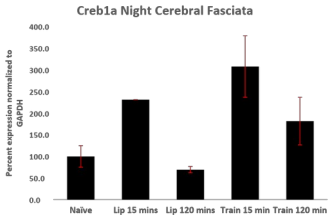
One of my roles in the lab involved purifying RNA samples extracted from animals. Using a nanodrop spectrophotometer, I assessed sample purity, with a 260/280 nm ratio of approximately 1.8 generally indicating a pure RNA sample. If the sample generated a ratio below 1, the sample was cleaned using a modified lithium chloride precipitation protocol with ethanol. The box plot in Figure 1 shows the effectiveness of the RNA cleaning process on about 40 samples that I cleaned. The left box has higher variation and a larger range, indicating that the RNA purity was often low before cleaning. The right box, in contrast, has a smaller range, indicating that the RNA became more pure after cleaning.



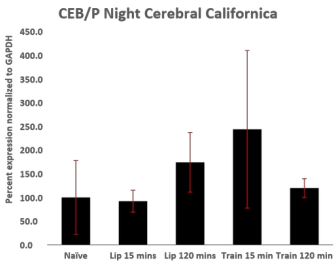
(a)



(b)



(c)



(d)

Depicted above are several graphs generated from quantitative PCR on the genes *Creb1a* and *C/EBP* from the buccal and cerebral ganglia of both the *Aplysia californica* and and *Aplysia fasciata* under various conditions. It is important to note that *A. californica* are diurnally active, and training them during the day is effective in producing memory, but training at night is ineffective. By contrast, *A. fasciata* are nocturnally active, and only training at night is effective. While these results are preliminary, they indicate that gene expression may be lower in the diurnal slugs when trained during their non-waking hours. Interestingly, it seems that 15 minutes after training, *C/EBP* and *Creb1a* are highly expressed but after 120 minutes, the gene expression is reduced (information from graph 2a regarding 120 min training is missing and still needs to be clarified). This could suggest that there is an initial spike in gene expression that is insufficient to stimulate consolidation, possibly due to blockage by a natural memory inhibitor. Furthermore, when nocturnal slugs were trained at night, there was a positive correlation with high gene expression of *Creb1a* and *C/EBP*, as seen in the graphs, supporting the correlation between expression of these genes and memory consolidation.

While these findings are preliminary, they suggest a significant role for *C/EBP* and *Creb1*a in memory consolidation and indicate differences in gene expression between nocturnal *Aplysia fasciata* and diurnal *Aplysia californica*. This research could provide us with a greater understanding of learning and memory processes as well as enhance our molecular knowledge of *Aplysia*, specifically *fasciata*, (as almost all *Aplysia* work at the genetic level has been thus far on c*alifornica)*. Down the line it could possibly have practical pharmacological implications on memory-related diseases, such as PTSD.

[1] Kandel, Eric R. *The Physical Basis of Memory*. Nobel Prize Lecture, 2000. Nobel Prize, https://www.nobelprize.org/uploads/2018/06/kandel-lecture.pdf.

[2] Roi Levy, David Levitan, Abraham J Susswein (2016). *New learning while consolidating memory during sleep is actively blocked by a protein synthesis dependent process* eLife 5:e17769. https://elifesciences.org/articles/17769

Analyzing Alzheimer's Disease: Measuring Amyloid-Beta in the Brains of Transgenic 5xFAD Mouse Models

Esther Nahon
Advised under Prof. Eitan Okun and Master's student Dolev Mikalovitch

Alzheimer's Disease (AD) is one of the most prevalent neurodegenerative diseases in the world associated with memory loss and dementia and unfortunately lacks effective treatment. AD is commonly characterized by the aggregation of Amyloid-beta (Aβ) plaques and neurofibrillary tangles (NFTs) in the neural brain tissue. Aβ is a peptide that is cleaved from the Amyloid precursor protein (APP), a protein encoded by the APP gene located on chromosome 21. Those exhibiting Trisomy 21, also known as Down syndrome (DS), express an extra copy of the APP gene which makes them more susceptible to developing AD at an early age. Prior studies have measured the presence of Aβ in the Dp1Tyb mouse model for Down syndrome which does not exhibit Aβ in the form of plaques. Focused research on different isoforms of Aβ in both transgenic and wild type (WT) mice proves to be a good method for studying AD in DS mouse models. The current research is similarly focused on measuring these Aβ isoforms in a different transgenic mouse strain, called 5xFAD, as they over-express the

human variant of mutant APP, thus making them another model for studying AD.

The purpose of this research is to calibrate an effective, affordable, and accessible protocol to measure and compare the levels of human Aβ (hAβ) isoforms Aβ1-40 and Aβ1-42 in the brains of 5xFAD mice that have been vaccinated for AD against WT and non-vaccinated 5xFAD mice.

A sandwich enzyme-linked immunosorbent assay (sELISA) was used to test which antibodies best detected the presence of Aβ in the 5xFAD mouse model. Multiple antibody and protein calibrations were performed on soluble and insoluble samples taken primarily from the cortex and hippocampus of 5xFAD mice upon which this method was used. The EPR23 capture and 6E10 detection antibodies were used to measure the Aβ1-40 isoform in the cortex and hippocampus while the 6E10 capture and 1-11-3 detection antibodies were used to measure the Aβ1-42 isoform in the af brain regions. Standard curves were created for reference during each calibration using a serial dilution method (Figure 1). Finally, spectrophotometry was used to measure the optical density (OD) of the samples, allowing for the calculation of the total protein, and ultimately Aβ concentration present in the sample.

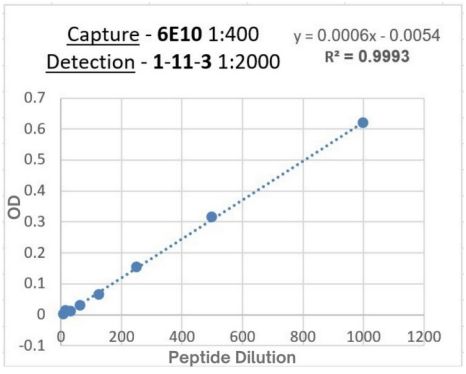


Figure 1: The ready-to-use dilution obtained from the antibodies calibration for the Aβ1-42 isoform

The sELISA technique can be used to test the effectiveness of an AD vaccine. In this study, six groups of mice including males, females, WT, 5xFAD, vaccinated for AD and HBS act as the subjects. The 5xFAD vaccinated group was injected with a vaccine for AD which is meant to reduce Aβ production. Following vaccine administration, cognitive and behavioral tests were performed on the mice in order to observe differences in their brain plasticity. Perfusion was then performed, upon which the mouse's brain was extracted, preserved, sliced, and stained to be further analyzed for its Aβ component. Though no results were consistent across the board, most improvement was found in the cognitive areas of female mice treated with the vaccine compared to the control group.

After perfecting the protocol, this sELISA method will be used to measure the Aβ levels in the experimental 5xFAD mice from the AD vaccine study. This will allow for comparisons to be made between the two isoforms in the 5xFAD mouse model as performed in previous studies with Dp1Tyb mice (Figure 2). Largely speaking, implementation of this method will help us better analyze the vaccine's effect against AD in this model.

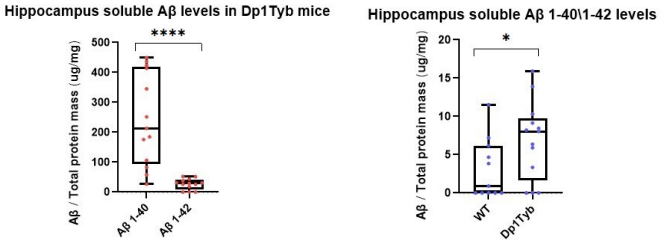


Figure 2. Statistical analysis previously performed on transgenic Dp1Tyb mice using a similar sELISA protocol to measure Aβ. Left: Overall comparison of the different Aβ isoform levels in the Dp1Tyb mice. Right: The ratio of Aβ1-40 and Aβ1-42 isoforms present in the hippocampus of Dp1Tyb mice compared to WT.

ENGINEERING



David Benhamu, Yishai Kohn, Tuvya Macklin, Betzalel Cohen, Jordan Paraboschi, Blimy Dresdner, Hannah Bialik, Hannah Zahler, Perie Hoffman

BiCCA Adaption of Canonical Correlation Analysis

Jordan Paraboschi
Advised under Dr. Ofir Lindenbaum, PhD Candidate Ran Eisenberg

Background

One common way to simplify high dimensional data analysis is to reduce the dimensionality of the data by projecting it onto the dimensions that are the most explanatory. This kind of technique can also be performed across different media using a particular method called Canonical Correlation Analysis, or CCA. CCA finds the dimensions in two different sets of data — as potentially disparate as, say, video and audio data — that have the largest correlation. The BiCCA algorithm is introduced to solve unique problems that may arise in applications of this method to RNA and ATAC sequencing found in computational genetics.

One core assumption of CCA is that the rows of the two data matrices correspond to the same label. This is called row or sample alignment. While modern RNA and ATAC sequencing technologies can characterize genetic information at the cellular level, there is no existent way to match the original cells that

potentially millions of randomly selected RNA and ATAC samples were sampled from, thus violating CCA's assumptions. The objective of this project is to capture the applicability and effectiveness of the BiCCA algorithm when this assumption is violated using the MNIST dataset, a dataset popularly used by academics in machine learning and data science. A sub-goal of this project was also to re-implement the BiCCA algorithm in Python instead of R.

Methodology

The results obtained in this study were all collected using the MNIST dataset. The MNIST dataset is a dataset of 60,000 training images and 10,000 testing images, all of which depict the digits 0-9 drawn by hand with two different noise regimes. This study only analyzed sets of 600-900 images for computational ease, although there is no reason to believe that this would compromise generalizability. All notions of error are evaluated using a five neighbor nearest-neighbors algorithm (5-NN). This algorithm makes predictions on the label of an unknown data point using a majority vote from the five nearest known data points. This technique is a measure of how the data points cluster. If the BiCCA algorithm is effective, images that represent the same digit in some dataset X should be projected near the same digit in dataset Y in the common subspace found by BiCCA.

Analysis and Results

Instead of performing a normal CCA operation, which assumes row-alignment, BiCCA creates a third matrix Z, which is row-wise correlated with X and column-wise correlated with Y. Z can be used to project X into a column-space correlated with Y, where row permutations are irrelevant. In the simulations shown in Figure 1 and Figure 2, there are three labels: images of either the digit 9,6 or 4. The error rate displayed in these graphs represents the data points in X that were incorrectly predicted using a 5-NN algorithm trained on Y in the respective subspace. The baseline is how well the labels of Y predict the labels of X without any projection. The two simulations shown in the figures each involved different percentages of row-misalignment (66% and 100%). As can be seen, the BiCCA algorithm performs close to the ideal "No Projection" error rate even despite row permutations, whereas the CCA projection performs increasingly worse.

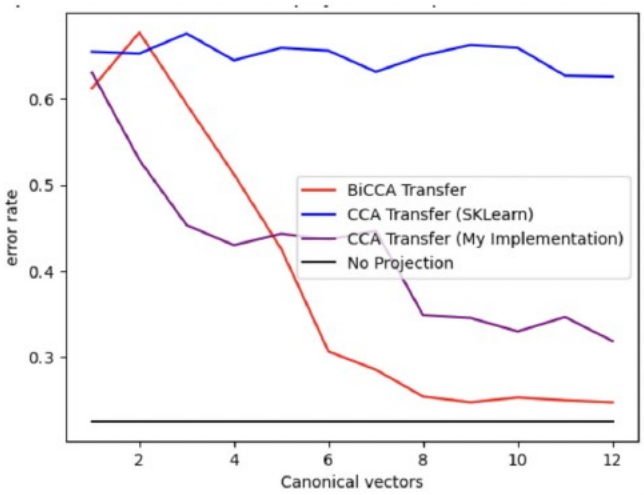


Figure 1: Labels 9,6,4 evaluated with 5-NN, 66% misalignment

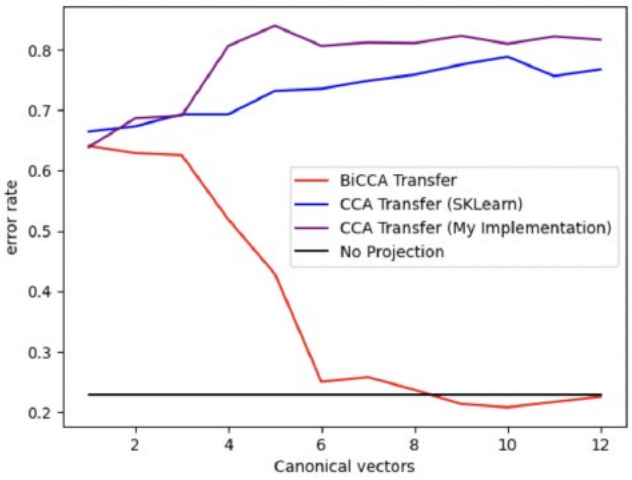


Figure 2: Labels 6,9,4 evaluated with 5-NN, 100% misalignment

As explained earlier, the real utility of CCA is in correlating features in different feature spaces. After all, if X and Y cluster just as well in the subspace as the general space, there may not

be a good reason to employ any projection regime. The plot in Figure 4 depicts a simulation with two labels (9 and 6) and 100% row misalignment, but this time the data matrices lie in different feature spaces. This simulation captures a novelty of BiCCA. In Figure 4, The image labeled X is an example from the X data matrix and the image labeled Y is an example from the Y data matrix. Notice that they each form a different half of the number 6. To avoid the sample-alignment assumption in CCA, BiCCA creates a third data matrix Z that represents the projection of the samples of X onto the feature space of Y, which in this case is the other half of the number. The error rate represents how well these projections cluster with where they should. As can be seen, BiCCA still finds a maintains an error rate remarkably close to the "No Projection" error rate in Figure 2, where the data in Figure 2 are in the same feature space.

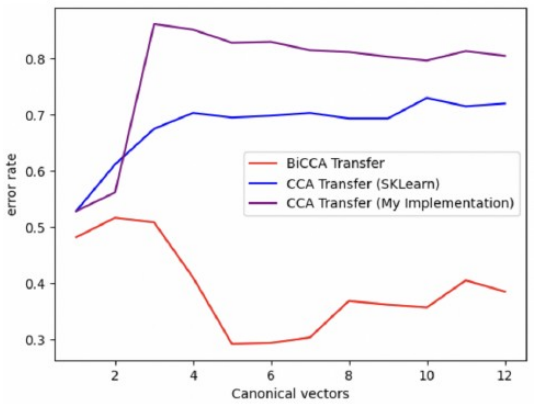


Figure 3 - labels 6 and 9 evaluated with 5-NN, 100% misalignment

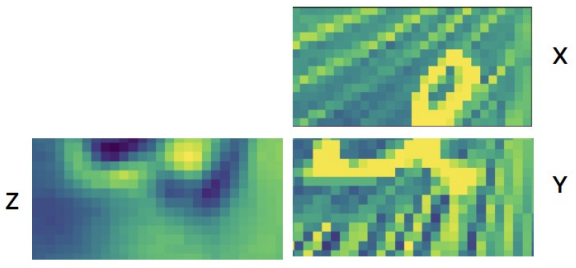


Figure 4 - examples from X, Y and Z matrices

These simulations raise a number of important questions: The BiCCA algorithm requires many initialization parameters, how can these parameters be chosen optimally? The MNIST dataset is a relatively simple dataset, how do these results generalize to more complex datasets? How effectively can BiCCA be used to recover the inital permutation of the misalignment? Answering these questions will be critical to understanding fully in what circumstances the BiCCA algorithm can be deployed and how it can be used most effectively. Despite these remaining questions, this project has succeeded in its implementation of the BiCCA in generalizing its results on a more easily interpretable dataset. Hopefully this project will provide a foundation for further research into the BiCCA algorithm.

Connecting and Evaluating the Reasoning Engine on a Repository of Real-world Boolean Network Models

Blimy Dresdner and Hannah Zahler
Advised under Prof. Hillel Kugler, PhD student Eitan Tannenbaum

For many years, biologists have explored gene interactions to gain a deeper understanding of the dynamic behavior of biological systems. Examining these interactions provides valuable insights into an organism's traits, inheritance patterns, potential mutations, and how system behavior changes under varying inputs or genetic modifications. However, as biological systems become more complex, the vast number of gene interactions makes manual study increasingly challenging. Therefore, there is an increasing need for advanced computational tools to analyze them. To address this complexity, scientists have turned to computational biology and advanced tools like Boolean Network models (Figure 1). These models are particularly effective for thoroughly analyzing gene regulatory networks, as they

clarify the logical relationships between genes[1]. Boolean Network models are used to visually represent the gene regulatory network. In these models, nodes represent genes, with directed edges indicating the interactions between them, and these genes can either be active (1) or inactive (0). The gene interactions are represented by directed edges between nodes. These edges are either positive, representing an activating relationship, or negative, representing an inhibiting relationship. This visualization captures the complex interplay between genes, providing a comprehensive overview of the network's behavior under various conditions.

This study aims to develop a comprehensive method for converting AEON and BNet files into REIN files, facilitating the use of reasoning engines. Reasoning engines are powerful tools that can automatically synthesize consistent Boolean Networks, offering a streamlined way to explore and experiment with the behavior of these systems. By converting and integrating these models, scientists can run simulations and design experiments to validate the networks. This integrated approach not only enhances our understanding of the dynamic behavior of biological networks but also provides a more efficient process for analyzing and experimenting with these models. The goal of our research is to develop a code to connect the reasoning engine to a cutting-edge repository of Boolean Networks, paving the way to evaluate the synthesis algorithms in the reasoning engine on a wide set of realistic biological models. This code will take BNET files of various models from the biodivine-boolean-models repository on GitHub and run them through bioLQM to determine their fixpoints[3]. Using these fixpoints, along with AEON files from the same models on GitHub, we will develop code to convert this information into REIN files (Figure 2). These REIN files can then be processed by the reasoning engine to verify the set of constraints and specifications for each model, ensuring they are satisfiable by checking if solutions can be found. To further improve the reasoning engine, we aim to automatically connect it to a state-of-the-art repository of Boolean Networks, opening the way to evaluate the synthesis algorithms in the reasoning engine on a wide set of realistic biological models.

We began by developing a code to convert the downloaded BNET file into a format compatible with bioLQM[3]. This conversion enabled bioLQM to process the file and compute the fixed points, which were then saved into an output file. Subsequently, we integrated this output file with the AEON file obtained from GitHub. Our code was then used to convert these combined files into a new format that is fully compatible with the reasoning engine. This streamlined approach not only facilitated the transition from BNET to REIN files, but also enabled the efficient analysis of Boolean models and the generation of experiments for further reasoning and validation. This file is then able to be put into the reasoning engine which automates the analysis and testing of this network. This converted REIN file is processed, and then a Constrained Abstract Boolean Network (cABN) diagram (Figure 3) of the model is created that is consistent with the experimental observations. Based on these experimental observations, the reasoning engine will determine whether solutions were found. This research has successfully developed a comprehensive method for converting AEON and BNET files to REIN files, facilitating the analysis of Boolean models and the generation of experiments for reasoning engines. The approach greatly simplifies the process of converting and analyzing biological network models, making it more accessible for researchers to study gene regulatory networks.

Furthermore, by obtaining fixed points from simulations, we provide valuable insights into the stable states of gene regulatory networks, which can inform further biological experimentation and validation. Ultimately, this approach connects computational modeling with real biological research, helping us better understand complex systems and opening up new opportunities for research and discoveries in systems biology.

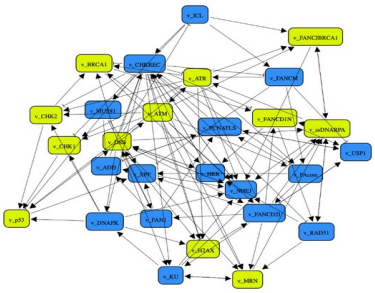


Figure 1: Boolean Network with genes as components. Arrows indicate interactions. Active genes are highlighted in yellow, while inactive genes are shown in blue.

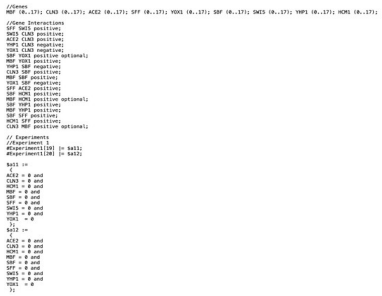


Figure 2: REIN file representation of model after reading in the files from the repository and running them through the code. This file contains all definite and optional gene interactions, and constraints defined as individual experiments.

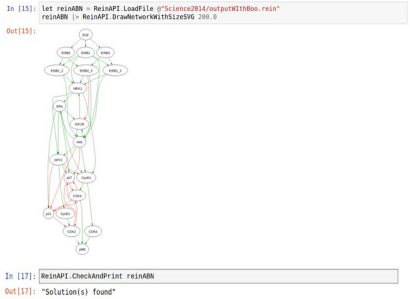


Figure 3: Constrained Abstract Boolean Network (cABN) which is output after the converted RE:IN file is entered into the reasoning engine and determined solutions found.

[1] B. Yordanov, S-J Dunn, H. Kugler, A. Smith, G. Martello and S. Emmott.
A method to identify and analyze biological programs through automated reasoning. NPJ systems biology and applications, 2,1–16, 2016.

[2] B. Yordanov , S-J Dunn, C. Gravill , H. Arora, H. Kugler and Christoph M. Wintersteiger. An SMT-based Framework for Reasoning about Discrete Biological Models. Journal of Computational Biology, 30(9), 1046–1058, 2023.

[3] S. Pastva, D. Safranek, N. Benes, L. Brim and T. Henzinger. Repository of logically consistent real-world Boolean Network models. bioRxiv 2023.06.12.544361

Measuring Blood Pressure Using Light

Betzalel Cohen
Advised under Dr. Hamootal Duadi, Prof. Dror Fixler, PhD student Michal Katan

Background

Measuring blood pressure (BP) is a crucial component of monitoring cardiovascular health, typically performed using traditional cuff-based methods. However, these methods can be cumbersome, uncomfortable, and inconvenient for continuous monitoring. Photoplethysmography (PPG), a non-invasive optical technique that measures blood volume changes in the microvascular bed of tissue, has gained attention as a potential alternative for BP measurement. PPG sensors are often embedded in wearable devices, offering a convenient means for continuous, real-time monitoring of physiological parameters. Despite its potential, the accuracy and reliability of BP measurements derived from PPG signals remain a significant challenge due to the complex relationship between blood pressure and PPG waveform features.

Problem

The primary problem addressed in this study is the development of a reliable method for

estimating blood pressure using PPG signals, with the goal of achieving accuracy comparable to traditional cuff-based measurements. The non-linear and dynamic nature of the cardiovascular system, along with the influence of various physiological and environmental factors on PPG signals, complicates the estimation process. Traditional machine learning techniques have shown promise, but they often require large datasets and extensive feature engineering. Additionally, these models may struggle with generalization across different populations or varying physiological conditions. To address these challenges, this study utilizes a sensor that can more accurately capture these PPG signals by measuring the absorption rays at a specific angle known as the Iso-path Length point (IPL) to enhance the accuracy of BP estimation from PPG signals. The key objectives are to:

1. Develop a sensor at this IPL point that captures the relevant characteristics of PPG signals for BP estimation.
2. Analyze this data to effectively map these features to BP values, accounting for the variabilities in the data.
3. Validate the proposed method against a benchmark dataset and compare its performance with existing methods.

Techniques Used

The proposed approach consists of several key components:

1. Preprocessing: The PPG signals are preprocessed to remove noise and artifacts, which can otherwise obscure relevant features. This step includes filtering, normalization, and segmentation of the signals.
2. Feature Extraction: A combination of traditional signal processing techniques (e.g., time-domain and frequency-domain analysis) and focusing on key points within the PPG signal (FIGURE 1) is used to extract features from the PPG signals to compare. The goal is to capture both local and global characteristics of the waveform that are indicative of blood pressure changes.

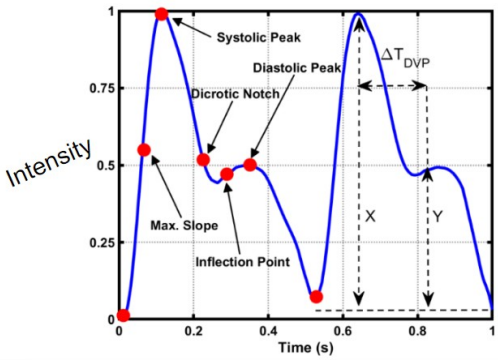


Figure 1: Filtered Photoplethysmography Signal Over Time (Example)

3. Validation: The model is validated using a separate test dataset, in which we measured 50 people, and its performance is evaluated in terms of accuracy and correlation with ground-truth BP measurements. Additionally, the metric's robustness is tested across different subsets of the population.

Results

Preliminary results indicate that the proposed method achieves promising accuracy in BP estimation, with a mean absolute error comparable to that of traditional cuff-based measurements and previous studies.

The validation on the test dataset shows a small correlation between the predicted and actual BP values, with minimal deviation across different population groups. The robustness tests also reveal that the model maintains its accuracy under varying conditions, such as changes in body position, talking, and skin tone.

Conclusion

This study presents a novel approach to blood pressure estimation using PPG signals, combining advanced signal processing techniques with a novel new sensor technique. The results suggest that the proposed method can achieve accuracy levels on par with traditional cuff-based measurements, while offering the added benefits of continuous, non-invasive monitoring. The analysis's ability to generalize across different populations and conditions underscores its potential for widespread application in wearable health monitoring devices.

Future work will focus on further refining the model, particularly in terms of creating deep learning models to enable real-time processing on low-power devices. Additionally, expanding the dataset to include a more diverse range of subjects and conditions will be crucial for improving the model's robustness and generalizability. The goal is to develop a reliable, user-friendly solution for continuous blood pressure monitoring that can be seamlessly integrated into everyday life and doctors' offices everywhere, providing valuable insights into cardiovascular health and aiding in the early detection of hypertension and other related conditions.

Computational Studies of PbSe Metasurface for Mid-Infrared Metaphotonics

Yishai Kohn
Advised under Prof. Tomer Lewi, Post-Doc. Researcher Sukanta Nandi

Introduction

Mie Theory was developed in 1908 by Gustav Mie, solving Maxwell's equations for the interaction of electromagnetic waves with an homogeneous "subwavelength" sphere of comparable size[1]. Within Mie theory, the multipole decomposition of particles emerges as a notable factor. As frequencies at which the material composition is resonant (the fundamental resonance can be roughly estimated using the equation λ=n*d, where "n" is the refractive index of the material and "d" is the diameter of the spherical subwavelength resonator) with the incident wave, the electromagnetic scattering is observed to mimic the fields of electric and magnetic poles[2] (Fig. 1 - dipoles, quadrupoles, octopoles etc). The contributions of these "harmonics" vary by material (i.e. index values) and can result in distinct electromagnetic properties like that of the Kerker effect. Mie theory has many applications - particularly in the field of metamaterial/metasurfaces, where the utilization of subwavelength structures in the order of nanometers in an array is a critical component in the creation of artificial material structures that offer tailored light-matter interactions not observed in natural materials (e.g negative-index).

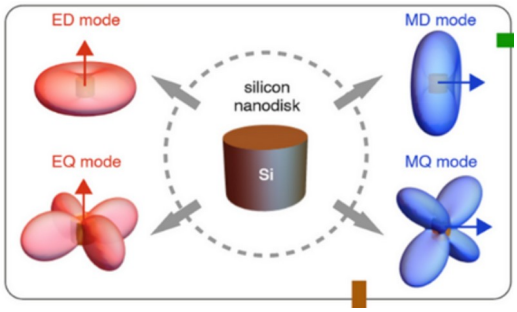


Figure 1: Resonant Modes for a Silicon Resonator

These metasurfaces have a wide range of applications from single photon detections to Bio-sensing. While Mie theory is strictly solved only for the homogenous sphere, metamaterials utilize alternative unit cell shapes for both ease of fabrication and modulation of scattering patterns.

Methodology

Finite-Difference Time-Domain (FDTD) is a method utilized to simulate light-particle interactions for both single unit components and complex devices in 1-3 dimensions. FDTD solves Maxwell's equations in the time domain and iterates these solutions over a mesh of variable precision. The mesh discretizes a given region of simulated space as Yee Cells (Developed by Kane S. Yee in 1966; Ref. 3), within which the electric and magnetic vectors are staggered, allowing simple computation of both fields. The respective field values are computed and averaged with those of their related cells at the midpoint using the finite difference method and iterated in every cell as the simulation evolves in time. Lumerical's FDTD software utilizes Yee Cells and allows one to choose from a variety of materials, source types, mesh precisions, monitors, etc. The commercially licensed Lumerical's FDTD program can be used to display electric and magnetic field profiles, in addition to many other visual and numerical simulation results.

FDTD is highly applicable in the field of nanophotonics, and has been utilized in a variety of technological fields. The ability to simulate electromagnetic wave/particle interactions is critical when developing many sensors, optical tools, photonic electronics, and lenses. In this work we computationally study the mid-infrared metaphotonic capabilities of chalcogenide (Lead Selenide, PbSe) metaatoms (single resonators) and metasurfaces (arrays of resonators).

Results

We studied multiple materials and resonator geometries, focusing mainly on full disk bodies at various aspect ratios (AR, diameter/height). We constructed contour plots to graph trends in scattering cross sections (SCS, the ratio of the intensity of radiant energy scattered to the incident irradiance), noting regressions of high resonance (Fig. 2 Silicon Resonator in Air: ARs 0.6 to 1.8). Particularizing our study to Lead Selenide (PbSe) we noted its improved performance due to its high refractive index (n≈4.7 and k=0, i.e. no material absorption) for the range of 5 to 10 microns. PbSe resonators had a greater light localization (quality factor, Q = fR /Δf, where fR is the resonant frequency and Δf is the full width at half maximum) and could be smaller than its Silicon counterpart to achieve the same resonant frequency.

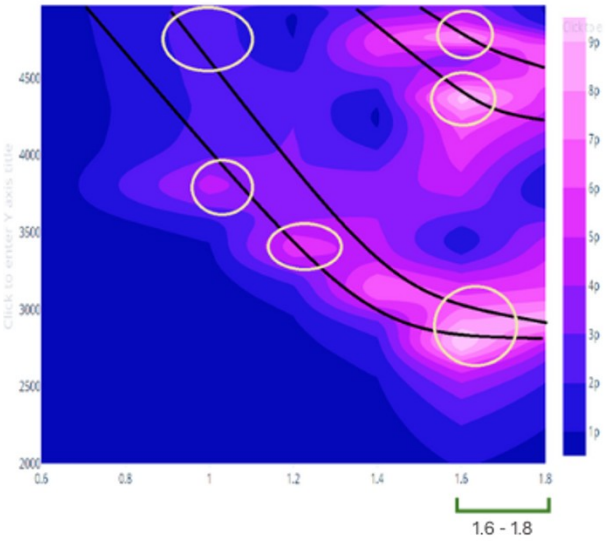


Figure 2: Contour Plot of SCS for SIlicon Resonator in Air

Optimizing the aspect ratios and periodicity (0.6 and 6600 nm), we constructed a metasurface utilizing these resonators and recorded the light reflection and transmission noting both a sharp second resonance and near-zero-reflective fundamental resonance in the transparent (within the bandgap) regime (Fig. 3). Focusing on this first result we postulated applications for index sensing (Fig. 4) simulating distinct environmental refractive indices. In addition to this result, the near-zero reflectivity achieved due to the fundamental mode presents applications for various anti-reflective coatings in the mid-IR electromagnetic spectrum.

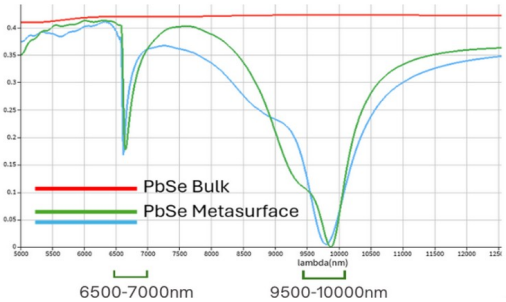


Figure 3: Mid-IR Reflectivity for PbSe Metasurface

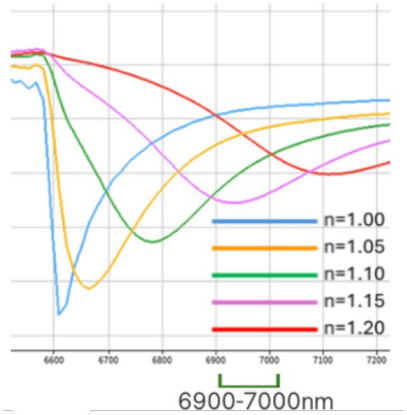


Figure 4: Second Resonance for PbSe Metasurface in Various Environmental Indices

Conclusions

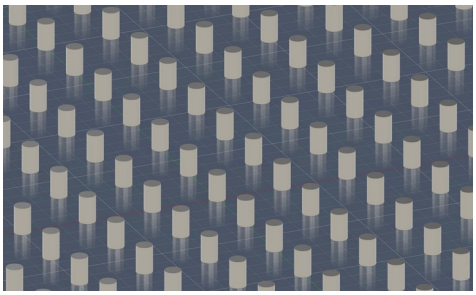
1) Finite Difference Time Domain is an effective tool for studying various light-matter interactions in nanophotonic systems.

2) Utilizing high index chalcogenides like PbSe to fabricate metasurfaces (Fig. 5) optimizes their metaphotonic capabilities. The diameter of a PbSe sphere resonator can be 28% smaller than its Silicon counterpart to achieve the same fundamental resonant frequency. In addition, the three modes of PbSe are sharper than those of Silicon with PbSe's fundamental mode Q value ~2.5 times that of Silicon. This trend comports with expected refractive index and Q value relationships.

3) The designed PbSe disk-resonator metasurfaces can be utilized for index sensing. As external n values change, the sharp second resonance (of Q value 66 in PbSe metasurface AR= 0.6) shifts with a sensing figure of merit ($S= \Delta\lambda/\Delta n$) in the order of magnitude 10^3. This leads to dynamic tunability of the metasurface which is useful for optic monitor and sensor applications.

4) The designed PbSe disk-resonator metasurfaces can act as anti-reflective coating. The reflection coefficients of both the AR= 0.5 and AR= 0.6 resonator surfaces drop to near-zero at ~9800 nm. This near-zero reflectivity, in contrast to the PbSe bulk reflectivity (~43%) at equivalent wavelengths, posits the efficacy of a metasurface coating that inhibits particular ranges of mid-IR frequencies. This result, combined with resonant tuning and multilayered metasurface lenses, can have practical applications for frequency inhibitors and various electromagnetic sensors.

5) The designed metasurface in this work thus provides a step forward towards the quest for miniaturization of nanophotonic components with enhanced tunable responses.


Figure 5: Simulated Optimized PbSe Metasurface

[1] Bohren, Craig F., and Donald R. Huffman. Absorption and Scattering of Light by Small Particles. Wiley : Wiley Online Library, 2007.

[2] Koshelev, Kirill, and Yuri Kivshar. "Dielectric resonant metaphotonics." Acs Photonics 8.1 (2020): 102-112.

[3] Yee, Kane. "Numerical solution of initial boundary value problems involving Maxwell's equations in isotropic media." IEEE Transactions on antennas and propagation 14.3 (1966): 302-307.

Multi-Agent Constraint CBS for Multi-Agent Path Finding

Tuvya Macklin
Advised under Dr. Dor Atzmon

Background

Multi-Agent Path Finding is a computer science problem where you are given a graph and a set of agents each of which have a start and goal location on the graph. The goal is to find the optimal set of paths for all the agents with no collisions. There are two basic types of collisions/conflicts. A vertex conflict occurs when two agents occupy the same vertex at the same time step. An edge conflict occurs when two agents cross an edge between the same timesteps in opposite directions. If a set of plans contains either of these conflicts, it is an invalid solution. A common measure of optimality, and the one used in our research, is called "Flowtime". This is evaluated by taking the sum of the time each agent takes to arrive at its goal from the initial time.

Disjoint CBS

Conflict-Based Search is a MAPF algorithm. It progressively develops a solution by identifying conflicts and placing constraints to resolve them. Initially, the algorithm uses A* to find paths for each agent, without considering the others. This is treated as the root node of the search tree. It then identifies a conflict to resolve. The conflict is resolved by constraining the vertex (or vertices in the case of an edge conflict) where it occurs. The constraint restricts one of the agents in the conflict from being at that vertex at that timestep. To explore all possible plans, two constraints are tried out, one for each agent in the conflict. They are applied to the scenario independently and form the first two child nodes. This is repeated for each child node recursively until a valid solution is found.

Disjoint CBS is a variant of CBS that alters the constraints used to resolve conflicts. It is intended to be an improvement on regular CBS. Instead of the two constraints restricting the two agents respectively, only one agent from the conflict is constrained. One constraint restricts the agent from being at that vertex at that timestep, and the other requires the agent to be there. This second constraint is called a "positive constraint". It effectively constrains all the agents since no other agent is allowed to be at that location since the chosen agent must be there.

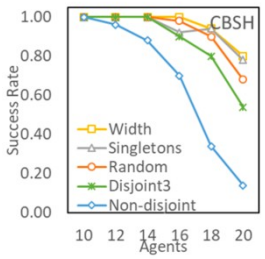


Figure 1: The success rate against number of agents for various versions of CBS. Note the performance increase of the disjoint versions.

We theorize that one of the reasons Disjoint CBS increases efficiency is because the positive constraints constrain many agents as opposed to just one. Based on this, we devised Multi-Agent Constraint CBS.

Multi-Agent Constraint CBS

Multi-Agent Constraint CBS (MAC-CBS) resolves conflicts in the following way. After identifying the conflict to resolve and separating the two agents involved, it takes the rest of the agents and splits them into two groups. Each group is assigned to one of the agents in the conflict and the constraints generated are applied to every agent in each group.

For example, if there are 10 agents and a conflict occurs between Agent 1 and Agent 2, the remaining agents (agents 3 through 10) will be divided into two groups. One group might include agents 3 through 6, and the other group might include agents 7 through 10. In one child node of the search tree, a constraint will be applied to agent 1 at the specific location and time step of the conflict, and the same constraint will also be applied to agents 3 through 6. In the other child node, a similar constraint will be applied to agent 2, along with the same constraint for agents 7 through 10.

Results

We ran tests using 50 grid-like instances. 25 were empty and 25 were randomly filled in 10% of the grid. Each instance was run 5 times with an increasing number of agents (4, 8, 12, 16, and 20). We ran each of these through regular CBS, Disjoint CBS, and MAC-CBS and gave each one minute to find a solution. The number of nodes expanded and whether the algorithm timed out were logged. See Figures 2 through 5 for the results. It seems that in terms of success rate, MAC-CBS performed poorly, even compared to regular CBS, but performed well compared to regular CBS in terms of nodes expanded. We think that the success rate is poor due to an issue with how the algorithm is implemented.

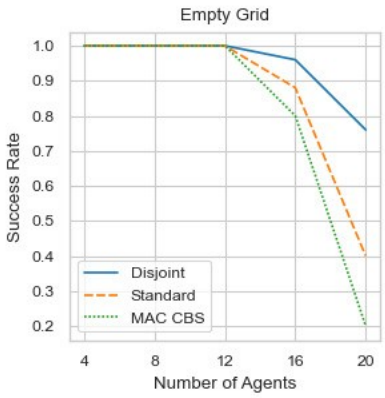


Figure 2

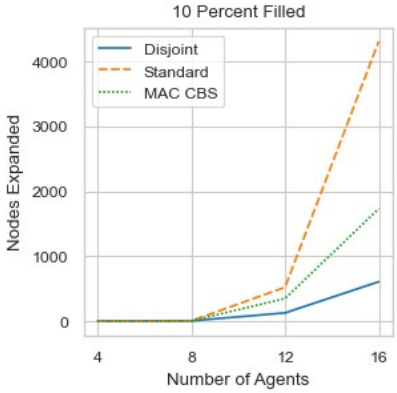


Figure 3

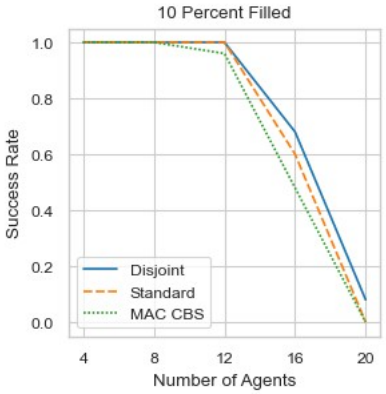


Figure 4

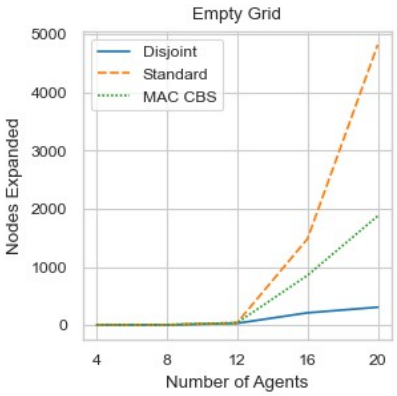


Figure 5

The empty map benchmarks were taken from the MovingAI benchmark library. They can be found at https://movingai.com/benchmarks/mapf.html

Conclusion

The analysis suggests that Multi-Agent Constraint CBS (MAC CBS) shows promise in standard Multi-Agent Path Finding (MAPF) due to its ability to reduce the number of nodes expanded during the search. This indicates potential for further efficiency improvements. Future work could focus on optimizing MAC CBS to enhance its performance even more. Additionally, further work includes applying MAC-CBS to the Robust MAPF and Large Agent MAPF variants. It is challenging to apply Disjoint CBS to these problems for reasons that are beyond the scope of this article, however, it should be simple to apply MAC-CBS here. We attempted to apply MAC-CBS to Robust MAPF but were not able to due to time constraints.

[1] Li, J., Harabor, D., Stuckey, P. J., Felner, A., Ma, H., & Koenig, S. (2021). Disjoint Splitting for Multi-Agent Path Finding with Conflict-Based Search. Proceedings of the International Conference on Automated Planning and Scheduling, 29(1), 279-283. https://doi.org/10.1609/icaps.v29i1.3487

[2] Stern, R., Sturtevant, N., Felner, A., Koenig, S., Ma, H., Walker, T., Li, J., Atzmon, D., Cohen, L., Kumar, T. K. S., Boyarski, E., & Bartak, R. (2019). Multi-Agent Pathfinding: Definitions, Variants, and Benchmarks. arXiv preprint arXiv:1906.08291.

[3] Atzmon, D., Stern, R., Felner, A., Wagner G., Barták, R., Zhou, N. F. (2018). Robust Multi-Agent Path Finding. International Symposium on Combinatorial Search (SoCS), 2-9

Exploring Single Photon Emitters in Hexagonal Boron Nitride for Quantum Applications

David Benhamu
Advised under Prof. Doron Naveh and Master's student Eyal Shoham

Single Photon Emitters (SPEs) play a pivotal role in the rapidly advancing fields of quantum communication, computation, and sensing. These unique quantum light sources, which emit photons one at a time, are particularly significant due to their potential applications in developing secure quantum networks, quantum computing systems, and highly sensitive quantum sensors. Among 2D materials, SPEs are especially abundant and easily detectable, attributed to both their perpendicular orientation relative to the material plane and the inherent defects commonly found in these materials. In this study, hexagonal boron nitride (hBN), a well-known 2D material, was extensively investigated for its potential to host SPEs, with a particular focus on identifying both individual and theoretically correlated emitters.

The primary objective of this study was twofold: First, to locate and characterize SPEs within hBN flakes, and second, to test the hypothesis of theoretically correlated SPEs[1]. A correlated SPE is a theoretical construct wherein two distinct SPEs are connected in such a way that they emit photons from different spatial points without overlapping, maintaining quantum coherence across distances. Although several measurements appeared to support the existence of this phenomenon, the results were inconclusive, requiring further experimentation and analysis.

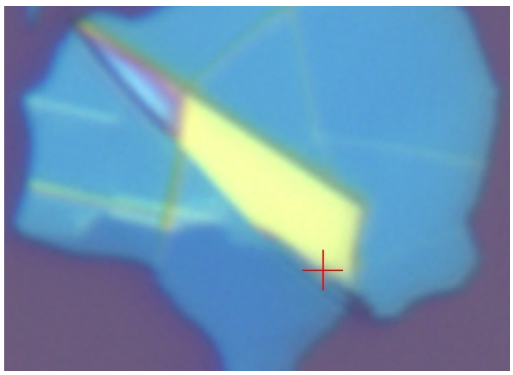


Figure 1: flake of hBN tested, red x marks one of the main points of interest.

Our methodology centered around the use of a confocal microscope paired with advanced spectroscopic tools to achieve high-resolution imaging and mapping of hBN flakes. After selecting a suitable flake of hBN, we generated photoluminescence (PL) maps by sequentially

exciting points on the flake using a laser and recording the resulting luminescence. These PL maps served as a critical tool for identifying potential locations of SPEs by highlighting areas of photon emission. Each promising location was then subjected to further scrutiny using a pair of avalanche photodiodes (APDs) connected to the system. By employing a beam splitter, the photon stream from the hBN flake was divided, with a 50% chance of a photon being detected by either APD.

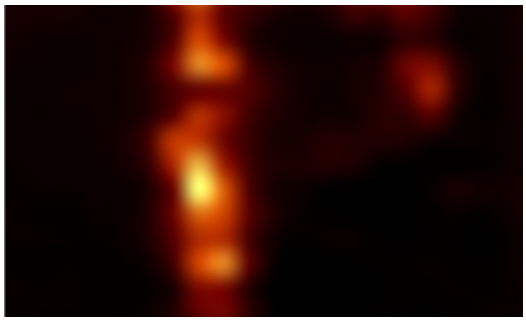


Figure 2: PL map of flake of hBN with point of interest circled

The use of APDs enabled us to perform time-correlated single-photon counting (TCSPC) measurements, which are critical for determining whether the photons emitted were indeed single and not part of a multiphoton event. Additionally, we observed photon "blinking," a phenomenon that indicates the intermittent excitation and de-excitation of SPEs. This behavior is crucial for identifying true SPEs, as it suggests the presence of a quantum emitter that alternates between "on" and "off" states, depending on the excitation conditions. Once sufficient blinking data was collected, we conducted correlation measurements to verify whether the observed emission pattern matched the behavior expected of a true SPE. These measurements involved fitting the data to the second-order correlation function, which allowed us to calculate the anti-bunching constant—a key indicator of single-photon emission.[2]

The results of this research hold promise for future advancements in quantum technologies. By pinpointing specific locations on hBN flakes where SPEs occur, we lay the groundwork for the development of quantum computing hardware, particularly in areas such as quantum memory storage.

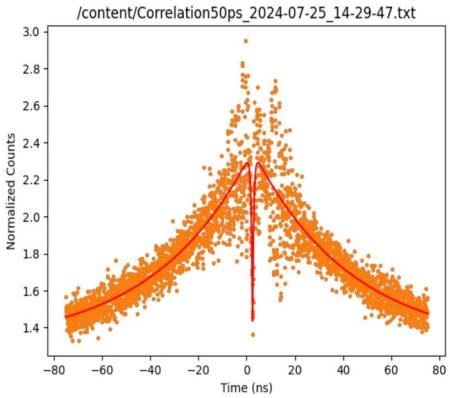


Figure 4: fit of the data to second order correlation and the discovery of an antibunching constant

Quantum computers rely on the precise control and manipulation of quantum states, and SPEs offer a robust mechanism for encoding quantum information at discrete points on a chip. Additionally, this work has broader implications for the field of superconductors, as SPEs can potentially be integrated into superconducting quantum circuits, further pushing the boundaries of what is possible in this cutting-edge field.

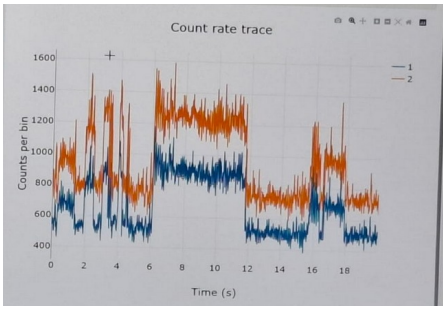


Figure 3: example of blinking in the APDs

Although the experiment did not conclusively demonstrate the presence of correlated SPEs, it provided valuable insights into the behavior of SPEs in hBN. Many new instances of anti-bunching were recorded, adding to the growing

body of knowledge about the types of defects in hBN that give rise to SPEs. These findings contribute to a deeper understanding of SPEs and their mechanisms, enhancing the potential for practical applications in future quantum technologies.

[1] M. Amazioug et al, Springer, 22, 301 (2023)

[2] S. Castelletto et al, Beilstein, 11, 740-769 (2020)

Detection of Atto 532 Fluorescent Dye Using a High Throughput Optical Modulation Biosensing System

Perie Hoffman
Advised under Professor Amos Danielli and Dr. Linoy Golani

In the early stages of a disease, biomarkers exist at very low concentrations and thus prove difficult to detect and analyze. Commercial immunoassays lack the combination of sensitivity and speed to detect low concentrations of biomolecules. Many immunoassays use magnetic beads but are not rapid or sensitive enough. These systems detect fluorescent signals from each individual bead, for instance BioPlex 2000 of Bio-Rad company, which makes them a timely process and they require sophisticated bulky equipment. The gold standard assay is typically ELISA (enzyme-linked immunosorbent assay) in which the fluorescent signal is obtained from the entire well. A high-throughput optical modulation bio sensing (ht-OMBi) system uses magnetic beads that capture fluorescently tagged molecules. This differs from ELISA in that the fluorescent signal is concentrated into the center by the magnetic beads and the sensitivity of the assay increases. This system aggregates the beads with a permanent magnet with a conical tip and is able to eliminate background noise by navigating a laser beam into the cluster of beads. The cluster of beads produces a greater fluorescent signal compared to that of an individual bead.

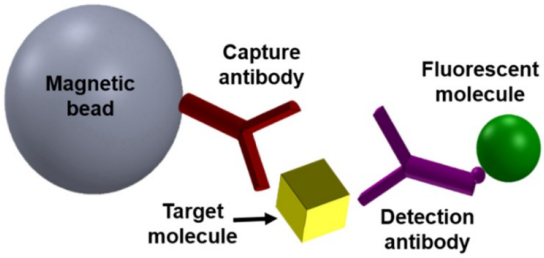


Figure 1: A typical fluorescent based sandwich assay. The assay consists of a magnetic bead coated with a capture antibody, a target protein, and a fluorescent lay labeled detection antibody[1]

Streptavidin-coupled magnetic beads were incubated with X10 dilution series of biotinylated Atto 532 fluorescent dye in order to find the limit of detection (LoD) of the assay. Photobleaching the magnetic beads prior to the experiment can increase the sensitivity of the assay. Three different experiments were carried out with beads photo-bleached for 18, 20, and 24 hours in order to determine the optimal bleaching time for the beads which will produce the most sensitive assay. Three more experiments were performed with non-bleached beads with different incubation times of 60 minutes, 30 minutes, and 5 minutes to determine whether the assay incubation time can be shortened. The OMBi system was able to detect various concentrations of Atto successfully in rapid time. Figure 3 displays a dose response using photo bleached beads for 18, 20, and 24 hours with 24 hours having the lowest LoD and therefore highest sensitivity. Figure 4 represents full dose response of Atto molecules in different incubation times of 60, 30 and 5 minutes. The LoD value was calculated for each experiment and was found to be 95.02 fM, 2.85 fM, and 39.47 fM respectively. The results demonstrate that there is no significant difference between 60- and 5-minutes assays, and the duration time of the assay can be shortened. Figure 2 illustrates how the OMBi system read and calculate the florescence signal: the florescent signal obtained from the background (B) subtracts from the signal obtained from the magnetic beads (A).

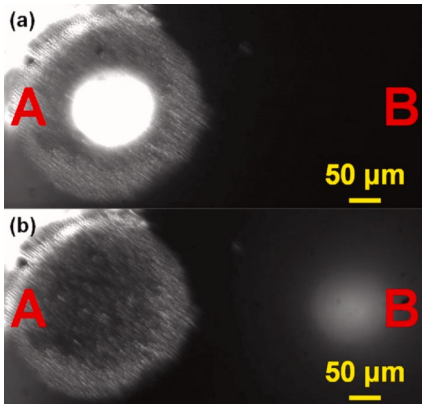


Figure 2: (a) The laser beam is steered from the left side of the well where the beads (A) are fixed. (b) the laser beam then moves to empty side of the well to detect the background signal (B).

During the photobleaching process, the magnetic beads are captured and illuminated by the laser for 2 minutes, and the arm is rotated to the middle of the cuvette for 30 seconds which allows the bead to sink. The arm then blocks the laser and grabs the beads on the other side for 10 seconds. The arm returns to the first position and illuminates the beads for 2 minutes and then the process is repeated. The rotation of the arm prevents the beads from quenching and declines their connectivity which reduces their auto fluorescence. The results indicate that 24 hours is the optimal bleaching time as seen by the lower LoD. Additionally, figure 4 displays that an incubation time of 5 minutes can be used since the LoD for 5-minute protocol is similar to that of the 30-minute protocol. This allows the experimental time to be shortened by x12.

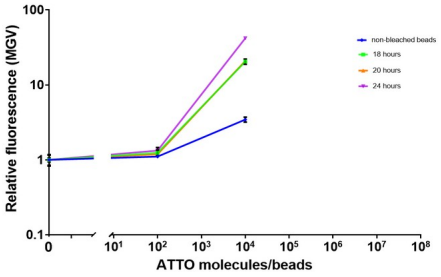


Figure 3: The relative fluorescence for three different dose response experiments. Each experiment was carried out with bead that were photobleached for various amounts of time.

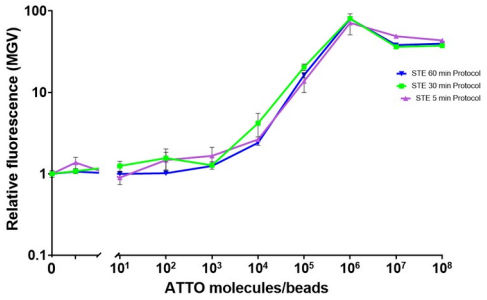


Figure 4: The combined standard error graphs for three biological repetitions. Each one was carried out with varying incubation times: 60 minutes, 30 minutes, and 5 minutes

Atto-OMBi based experiment was originally performed by non-bleached beads. We have demonstrated that using a bleached beads improves the sensitivity of the assay, specially 24 hours of photo-bleaching (Figure 3). Moreover, the Atto-OMBi based assay duration can be reduced to 5 minutes (Figure 4), enabling a rapid testing process. This shortened timeframe addresses the critical need for efficiency in immunoassays. These Atto experiments form the foundation for the development of a range of immunoassays designed to identify various pathogens and biomarkers.

[1] Roth, S., Margulis, M. & Danielli, A. Recent advances in rapid and highly sensitive detection of proteins and specific DNA sequences using a magnetic modulation biosensing system. Sensors 22, 4497 (2022).

[2] Burg, S et al. High throughput optical modulation biosensing for highly sensitive and rapid detection of biomarkers. Talanta 248, 123624 (2022).

Retention Time Tests on Gain Cell Embedded Dynamic Random Access Memory

Hannah Bialik
Advised under Prof. Adam Teman and PhD student Odem Harel

Computers require the ability to store information to harness user data effectively. Memory embedded in a microchip transmits

data more quickly and at lower power costs than external memories. Typically, Static Random Access Memory (SRAM) is utilized as embedded RAM because it is quick and reliable. However, with each bit-cell requiring 6 transistors, the storage density of SRAM is quite low. Embedded Dynamic Random Access Memory (eDRAM) macrocells have been developed as a higher density alternative to SRAM. The traditional version of eDRAM (1T-1C) is built with one single-access transistor and a special high-density 3D capacitor. The atypical manufacturing process of the high-density capacitor makes the 1T-1C eDram impractical in many situations. Gain Cell eDRAM (GC-eDRAM) is an alternative form of eDRAM with a basic bit-cell of 2-4 MOS transistors that can be easily integrated into a chip. As with all types of dynamic memory, data is stored as electric charge on a capacitor. Therefore, leakage currents compromise the stored data after some time, creating the necessity for periodic refresh operations to ensure that the data is not lost. The main downside of CG-eDRAM is its low retention time, which results in higher power consumption due to more frequent data refresh cycles. The Emerging Nanoscaled Integrated Circuit and Systems (EnICS) lab is developing methods of extending the refresh period - the length of time between refresh operations without losing the data - so that GC-eDRAM can be a more practical eDRAM option.

My main role at EnICS has been to test the data retention time of the GC-eDRAM on the Lupulus and Leo2 chips and to create graphical representations of that information. This involved writing, editing, and combining C and Python codes that interacted with the chips and a DC voltage supply. I also manually read and adjusted voltages on the research board that connects to the chip. In a most basic bit retention time test, a pattern of binary bits is written to an area of the chip in the form of an N-bit word. There is then an idle period after which the stored data is read. If there is a mismatch between what was previously written and what is read after the idle period, that means some bits failed. The retention time of a bit is the length of time it takes until the first time it fails, meaning that it flips (from 0 to 1 or vice versa) from the value that was originally written. However, since we can only write and read to a word at a time (32 bits in the case of the Leo2 and Lupulus), the read data must be parsed and bits that flipped for the first time in this reading must be identified. The retention time of these bits is equivalent to the idle period between writing and reading the word. After these calculations, if not all bits in the word have failed once, the word is written again with an increased idle period. Bits that fail for the first time in that round have a retention time of the increased idle period. This cycle continues until all bits in the word have failed at least once. The flowchart in Figure 1 illustrates this process.

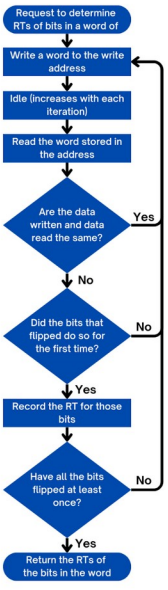


Figure 1: Flow Chart of Bit Retention Time (RT) Code.

Once this data was collected, I produced heatmaps like the one in Figure 2. Each square represents a gain cell which stores a bit of information. The Data Retention Time (DRT) of that bit is correlated to the square's color. The retention times in the figure are determined based on writing all ones to the bits, which is the worst-case scenario since ones have a lower retention time than zeros due to hardware design decisions. (This means writing 0xffffffff to each 32-bit word.)

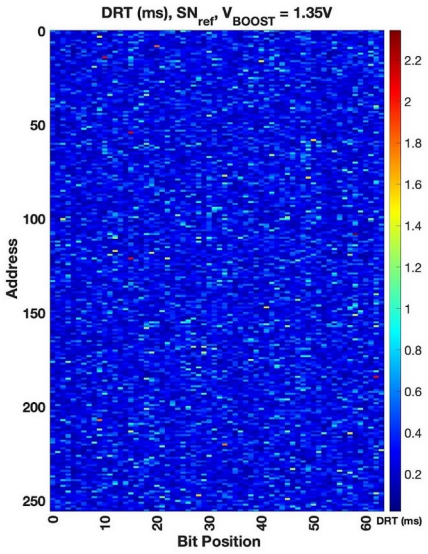


Figure 2: Heatmap of Data Retention Time of 2T Gain Cells in the Leo2 Chip.

While there is some variation in the retention time of cells in this sample, they are overall similar. This indicates relative uniformity among the transistors.

Inspired by concepts in Leo2 code (since for that chip I modified code that I had received), I set out to write the code to test the Lupulus retention time. Since it is tedious to change the input voltages with a screwdriver on the test board (and we wanted to test the retention time with various voltage inputs), we planned to hook the board up to an external DC power source and use the code to manipulate the input voltages. After some struggles accessing the correct location on the chip, my code worked and properly connected to the DC power supply (my code for that part was heavily based on a sample I received).

While the code seemed to be correctly reading the retention times, we received only about 99.96% accuracy per array (each array is 512 rows of 16 words each - 262144 bits altogether) at zero retention time (meaning we wrote and then read the data with almost no delay between these operations). This was surprising, as we had expected about 100% accuracy at such a low retention time. We did several trials and determined that it was not the same bits failing each time, but rather a random selection. We broadened our ranges of input voltages in hopes that perhaps an ideal combination of input voltages would lead to better results, but 99.96% accuracy was about the closest we could get.

Meanwhile, at ideal voltages about 7 bits per word were never failing no matter how long the idle period was. Testing the code using unideal voltages confirmed the code's accuracy via the never-failing bits (because in that situation all bits fail at some point), so we seemed to be facing hardware difficulty. Troubleshooting this difficulty helped me learn more about the inner workings of Lupulus gain cells and the test board used with it. I also gained a deeper understanding of the interaction amongst the transistors in the gain cell. However, we were unfortunately unsuccessful at pinpointing the problem and determining a solution during my time interning at the lab.

I am thankful for the opportunity to learn about gain cells, the Lupulus and Leo2 chips, and the behaviors of transistors. My time at EnICS also gave me a greater appreciation for the scope and capabilities of C, Python, and the interaction between software and hardware in memory chips.

LIFE SCIENCES



Alyssa Polonetsky, Tobi Katz, Hannah Mamet, Ella Nasser, Meira Strauss, Lilly Glickman

Genome Editing by CRISPR-Cas9 System for Modeling and Correction of Hematological Diseases

Tobi Katz
Advised under Prof. Ayal Hendel and PhD students Nimrod Ben Haim and Orli Knop

Background

CRISPR (clustered regularly interspaced short palindromic repeats) is an intrinsic feature of the adaptive immune system of prokaryotes. Bacteria utilize this system to protect themselves against foreign entities, and are able to reuse the same mechanisms of defense against any antigen. CRISPR mechanism is used for genome editing purposes while utilizing the Cas9 nuclease and a guide RNA (gRNA) to ensure specificity of edits (Fig. 1).

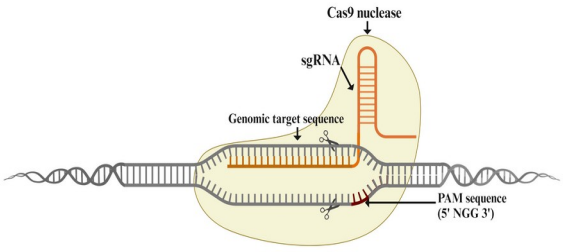


Figure 1: CRISPR-Cas9 System in alignment with the proper genomic target sequence, flanked by the PAM sequence

Here are three main steps in the CRISPR-Cas9 gene editing process. The gRNA first directs the Cas9 nuclease to the proper area of the genome near the Protospacer Adjacent Motif (PAM) sequence, which is denoted as NGG in our plasmid. Then, a complex travels along the genome until there is exact alignment between the spacer sequence on the gRNA and the base pairs on the DNA target. Third, the Cas9 nuclease cleaves the DNA forming a clean double stranded break (DSB.) The DSB then causes a cellular mechanism of repair to ensue, allowing editing to occur. There are two main pathways that can repair a DSB. The first is called the non-homologous end joining method (NHEJ). This occurs when small insertions or deletions (indels) are added to the genome to fix the breakage in the DNA. The other mechanism used is accomplished through homology-directed repair (HDR). This method involves using homology arms to insert desired portions into the gene through the use of plasmids. HDR can be induced over NHEJ by delivering DNA templates that can be inserted into the cut site, thus editing the gene[1]. This study focused on a gene that when mutated causes hematological diseases. These diseases are the first types which have been proposed to be cured using CRISPR gene editing since they are found on hematopoietic stem cells which can be isolated and engineered outside of the body. This study focused on a gene that when mutated causes hematological diseases. These diseases are the first types which have been proposed to be cured using CRISPR gene editing since they are found on hematopoietic stem cells which can be isolated and engineered outside of the body.

Methodology

The first step in this project was to design proper gRNAs which direct the Cas9 enzyme to the correct area of DNA for editing to take place. We designed gRNA to cut near the ATG sequence using bioinformatic tools. Five gRNAs were tested. There were two main laboratory experiments that were performed: plasmid cloning and guide competition. One method of delivering gRNA and the Cas9 protein into the cells is by using plasmids. The plasmid cloning process was performed for each of the five gRNA oligos to introduce them into the plasmid that constitutively expresses Cas9. During transformation, four colonies of plasmids were collected per gRNA plate. We also used a negative control containing no additional DNA, and a positive control using a plasmid that is known to work for the CRISPR-Cas9 system.

In plasmid cloning, we first digested the plasmid, allowing us to confirm that the plasmid was not contaminated. Next, we performed gel electrophoresis to separate the plasmid based on the bands, in order to ensure the cuts were in the proper location, and to purify the vector backbone. We then prepared the plasmid for oligo phosphorylation and annealing. The next step was placing the gRNAs and plasmids into the thermocycler to induce ligation. This step was followed with overnight ligation to increase the ligation efficiency. The result was a pool of plasmids. Following ligation, we set out to isolate individual plasmids to ensure they contained the proper DNA sequence matching with both the original plasmid, and gRNA insertion. We performed a transformation of the plasmids into bacteria, placing them on agar dishes. This process included plating, growing and screening the plasmids. Following transformation, we collected four colonies for each gRNA and underwent a plasmid miniprep to extract the plasmid. Next, we measured the concentration of DNA in each plasmid. Finally, we sequenced the plasmid and gRNA samples using Sanger sequencing. Sequencing results were compared to the original plasmid DNA and the specific sequences for each gRNA. Each of the gRNA

sequences were successfully inserted to the specific restriction site in the plasmid as can be shown on the chromatograms in the marked yellow regions.

Once the plasmid clones were formed, we then moved on to the second set of experiments. Hek293-Cas9 cells had been maintained throughout the length of the experiment, and were used to test the efficiency of the CRISPR-Cas9 gene editing system. We performed the editing of Hek293-Cas9 cells by each of the gRNA using electroporation (Fig. 2). DNA was extracted and a PCR was run with calibrated primers and sequenced with Sanger sequencing.

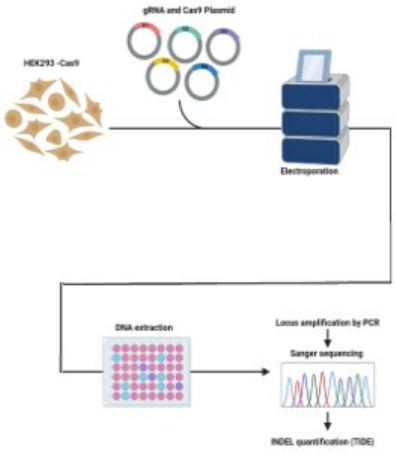


Figure 2: Process of Guide Competition Hek293-Cas9 cells are electroporated with sgRNA and a Cas9 Plasmid, DNA is extracted from the resultant and sent to sequencing to check efficiency of the CRISPR-Cas9 gene editing system.

Conclusion & Future Works

We successfully transfected the plasmids, as proven with Sanger sequencing comparisons. We performed a guide competition to test if the transfected plasmid correctly entered the Hek293-Cas9 cell which has been cultured during the duration of this project, and extracted DNA.

Our next steps include using the 'winning' guide to edit hematopoietic stem and progenitor cells (HSPCs) from healthy donors to knock out the gene. We will analyze how the deletion effects the phenotype, and subsequently design AAV6 donors to insert corrective cDNA into the gene locus.

I would like to thank Nimrod Ben Haim and Orli Knop for guiding me in this project, Nechama Kalter and Michael Rosenberg for their openness to questions, and of course Prof. Ayal Hendel for this opportunity.

[1] M.H. Porteus. "A New Class of Medicines through DNA Editing." The New England Journal of Medicine. 2019

HSV-1 Infection Expression on Cellular Factors

Hannah Mamet
Advised under Prof. Yaron Shav-Tal and PhD student Chaya Bohrer

Background

Herpes simplex virus type 1 (HSV-1) is spread by skin-to-skin contact and causes blistering sores around the mouth and lips. Currently, there is no known cure – once the virus enters the body, there is no way for it to exit; there are only medications and treatments an infected individual can be given to ease the symptoms. Like many other viruses, HSV-1 takes control of the host's cellular mechanisms to replicate itself. Specifically, we studied how HSV-1 takes control of the cell's splicing machinery.

mRNA is a single-stranded molecule from which proteins are made. DNA, composed of exons and introns, is transcribed by the enzyme RNA polymerase II in the nucleus. After transcription, this pre-mRNA strand undergoes modifications in the nucleus such as splicing. During splicing, introns, which do not code for an amino acid, are spliced out, while exons, which code for amino acids, are joined together. After splicing,

the now-mature mRNA is exported to the cytoplasm, where it undergoes translation to a protein. mRNA is degraded. This is the central dogma of gene expression.

The exon junction complex (EJC) is a family of proteins that are recruited to the splice site during pre-mRNA splicing and plays a key role in splicing. Y14 and Aly, are two proteins, or splicing factors, that are members of the EJC. Nuclear speckles are non-membranous bodies found in the nucleus and contain many splicing factors [1].

When HSV-1 enters the host cell, it expresses immediate early, early, and late genes[2]. US1, an immediate early gene, was studied under different conditions.

Since their genome is very condensed and small, viruses do not normally perform splicing. In complex organisms, like humans, splicing is crucial. Since the host needs to do splicing, the virus, in order to replicate its genes within the host cell, will also splice. Thus, in essence, the virus takes over the host cellular machinery, then silences the host machinery and afterwards, ejects itself and destroys the host cell in the process.

Methodology

The goal of this study was two-fold. First, the intention was to see the recruitment of cellular splicing factors Y14 and Aly to the transcription site of viral RNA. Second, was to examine the distribution pattern of Y14 and Aly under infection of HSV-1.

The techniques used to study these proteins – their localization and expression - were immunofluorescence and fluorescent in situ hybridization (RNA FISH). Immunofluorescence uses fluorophores to visualize the location of proteins. A fluorophore is conjugated to a secondary antibody which recognizes and binds to a primary antibody, and the primary antibody recognizes and reacts with the protein of interest. The wavelength emitted by the fluorophore allows us to visualize localization and expression of the protein of interest. RNA FISH is a molecular technique used to visualize RNA. In this technique, an RNA probe is labeled with a fluorophore and binds to the target RNA within the cell, allowing us to visualize gene expression.

Results

The distribution pattern of Y14 under varying conditions was studied using immunofluorescence. In Fig. 1, Y14 is stained green and DNA in the nucleus is stained blue. The first column across shows the untreated (mock) condition. Results from the magnified merge image show that Y14 is diffuse throughout the cell, but also found in higher concentrations in certain areas, which can be seen as darker green spots. These are the nuclear speckles, and as expected, splicing factors, like Y14, are found in greater concentrations there. In the row below, cells are treated with PLB, a splicing inhibitor. Under these conditions, there was not such a noticeable difference compared to the mock. It can be seen, however, that Y14 is slightly less diffuse throughout the cell. The third row shows the distribution pattern of Y14 2 hours post HSV-1 infection. A noticeable difference can be seen; Y14 is significantly less diffuse throughout the cell and is concentrated in the nuclear speckles, which are more noticeable now. The last row shows Y14 was tested under both conditions: PLB and 2 hours post infection. The magnified merge image shows that the combination of both these treatments caused the nuclear speckles to be more numerous and larger. We believe this is due to the combination of treatments having an inhibitory

effect on the recycling of splicing factors that are constantly moving throughout the cell.

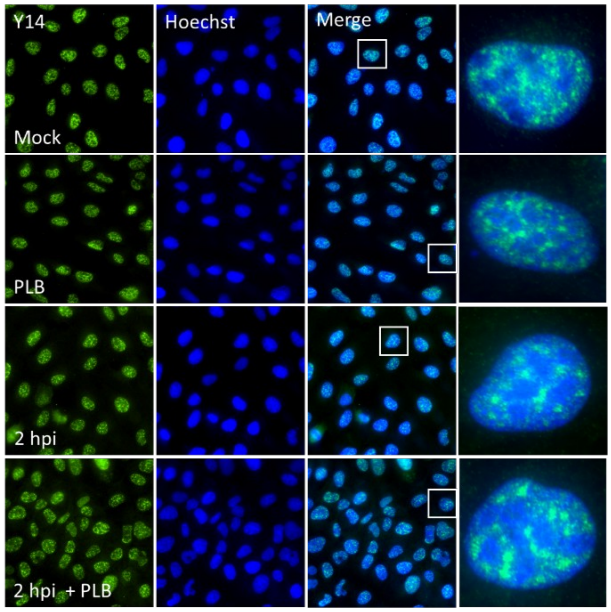


Figure 1: Results of immunofluorescence. Y14 condenses into foci under PLB and HSV-1.

The results of the immunofluorescence with Y14 prompted us to also study the distribution of another cellular splicing factor, Aly, and to study its expression with viral mRNA. To visualize the US1 viral mRNA and Aly, another splicing factor and protein that is part of the EJC, both immunofluorescence and RNA FISH were performed on the cells (Fig. 2). The first row shows the untreated conditions (mock); thus, no virus can be seen in the US1 viral mRNA image. The next channel is Aly, followed by a merged image. When comparing the mock results to the second row, 6 hpi, a drastic change can be seen. In the 6 hpi US1 viral mRNA channel, large concentrations of purple can be seen on the edge of the nucleus. The US1 mRNA is concentrated at the transcription site, as it works to express its own genes. Each purple dot at the cells' cytoplasm is a viral mRNA particle. The channel next to it, Aly, can be seen concentrated also in green pockets towards the periphery of the nucleus. The merge image demonstrates that US1 viral mRNA and Aly are colocalized to the transcription site. We can conclude from these results that Aly, an EJC protein, is being recruited to the transcription site in order to aid the virus in transcribing and expressing its own viral genes.

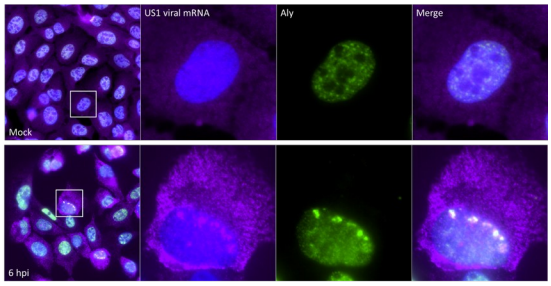


Figure 2: RNA FISH and Immunofluorescence show US1 and Aly are colocalized under HSV-1 infection.

Conclusion

Results from the immunofluorescence performed on Y14 demonstrate that when cells are treated in combination with PLB (a splicing inhibitor) and HSV-1, there is a much stronger effect on Y14 distribution. Results from the combined RNA FISH and immunofluorescence on Aly showed that the transcription site and Aly were colocalized, and therefore, we concluded that Aly is recruited to the viral transcription site.

[1] G. Faber, S. Nadav-Eliyahu, and Y. Shav-Tal, Journal of Cell Science, 135, 259594 (2022).

[2] F. Han et al, Journal of Virology, 92, 00536 (2018).

Survival Outcomes and Prognostic Factors in Acute Myeloid Leukemia Patients Through Rigorous Data Analysis

Lilly Glickman
Advised under Dr. Binyamin Knisbacher, Master's Student Eden Meidan

Background

Acute Myeloid Leukemia (AML) is a highly aggressive and lethal form of leukemia characterized by the rapid proliferation of abnormal white blood cells in the bone marrow.

These abnormal cells accumulate and interfere with the normal production of blood cells, leading to severe complications such as anemia, infection, and bleeding. AML predominantly affects adults, with a particularly high incidence in older individuals. The median age of diagnosis is approximately 68 years, and the disease is associated with a poor prognosis, particularly in older patients who may have comorbid conditions that limit the effectiveness of aggressive treatments. The heterogeneity of AML, both in terms of genetic mutations and clinical presentation, complicates treatment and management. Patients with AML exhibit a wide range of genetic abnormalities, which influence disease progression, response to treatment, and overall survival. As a result, there is a significant need for personalized treatment approaches that take into account the unique genetic and clinical characteristics of each patient. Identifying key prognostic factors that can predict survival outcomes and guide treatment decisions is critical in improving the management and outcomes of AML patients.

Goals and Methods

The objective of this research is to identify key prognostic factors in AML patients by conducting a comprehensive survival analysis. We aim to categorize patients into distinct subgroups based on their genetic and clinical attributes and to determine how these factors influence overall survival. By understanding the factors that drive disease progression and impact patient outcomes, we hope to contribute to the development of personalized treatment strategies that improve survival rates and quality of life for AML patients. To achieve this goal, we utilized a large dataset of RNA sequencing and clinical data from AML patients. The RNA sequencing data provided detailed information on gene expression levels, while the clinical data included variables such as patient age, white blood cell count, cytogenetic abnormalities, and treatment history. We began our analysis by preprocessing the RNA sequencing data to ensure consistency and accuracy across the dataset. This involved normalizing gene expression levels and filtering out low-quality data to focus on the most relevant and reliable information. Next, we applied a statistical method developed in our lab to identify gene expression outliers. These outliers were defined as genes whose expression levels significantly deviated from the median expression of the gene across the patient cohort. Specifically, we used a modified z-score to quantify the extent of deviation from the median, with a threshold set to identify genes that were aberrantly over- or under-expressed. Genes identified as outliers were further categorized as either oncogenes, tumor suppressor genes, or both, based on their known roles in cancer biology. This classification allowed us to explore the potential impact of these genes on AML pathogenesis and patient outcomes. To assess the clinical significance of the identified gene expression outliers, we conducted survival analysis using the Kaplan-Meier estimator. This statistical tool allowed us to generate survival curves for patients with different expression levels of the identified genes, providing insights into how these genetic anomalies influenced survival. We focused on comparing the survival outcomes of patients in the highest and lowest percentiles of gene expression, as well as those within the interquartile range (IQR). To ensure the robustness of our findings, we introduced statistical significance testing, including Fisher's exact test to assess enrichment within the subgroups. To control for multiple comparisons and reduce the likelihood of false positives, we applied a False Discovery Rate (FDR) threshold of <0.1. Additionally, we utilized several Python packages to streamline our data analysis and visualization. Pandas was used for managing the large, complex datasets, enabling efficient data

manipulation and aggregation. Matplotlib was employed to create detailed visualizations, including survival curves and heatmaps of gene expression data, which facilitated the interpretation of the results. The Lifelines package was essential for conducting the survival analysis, allowing us to generate Kaplan-Meier survival curves for each subgroup and perform Cox proportional hazards regression to assess the impact of multiple variables on survival simultaneously

Results

The results of our analysis revealed significant differences in survival outcomes based on the identified gene expression outliers. Patients with extreme deviations in gene expression—those in the top and bottom percentiles—exhibited markedly different survival rates compared to those with more typical expression levels, as represented by the IQR. This finding suggests that aberrant gene expression plays a critical role in influencing the aggressiveness of AML and patient prognosis. In particular, our analysis identified several genes that were consistently over or under-expressed in patients with poor survival outcomes. For example, genes such as BCOR, CEBPA, and FLT3 were found to be significantly over-expressed in patients with shorter survival times, indicating their potential role as drivers of AML progression (Fig. 1). Conversely, genes such as KIT, RUNX1, and STAG2 exhibited significant downregulation in these patients, suggesting that their reduced activity may contribute to the loss of tumor suppressive functions. The bar graph analysis further illustrated the distribution of these AML driver genes across the patient cohort. The x-axis of the graph listed various AML drivers, categorized as oncogenes, tumor suppressors, or both, while the y-axis represented the number of samples with significant up or down outliers. The graph highlighted the variability in expression patterns among these genes, with certain oncogenes like FLT3 showing a high frequency of upregulation, and tumor suppressor genes like STAG2 exhibiting notable downregulation. This variability underscores the complexity of AML and the need for targeted therapeutic strategies that address the specific genetic drivers of the disease in each patient. Furthermore, our enrichment analysis indicated that certain gene expression outliers were significantly associated with specific clinical features, such as cytogenetic abnormalities and white blood cell count. For example, patients with high expression levels of FLT3 were more likely to have adverse cytogenetic profiles, which are known to be associated with poor prognosis in AML. This finding suggests that FLT3 expression could serve as a biomarker for identifying high-risk patients who may benefit from more aggressive treatment strategies. Overall, our results demonstrate the importance of considering multiple factors in predicting survival outcomes in AML patients. By integrating genetic, molecular, and clinical data, we were able to provide a more comprehensive assessment of patient risk, which could inform treatment decisions and improve patient outcomes.



Figure 1: AML driver genes identified as outliers with a modified z-score greater than 2

Conclusion

In conclusion, this study provides valuable insights into the prognostic factors influencing survival in AML patients. Through a rigorous analysis of RNA sequencing and clinical data, we identified key gene expression outliers that correlate with patient outcomes and proposed

potential targets for therapeutic intervention. Our findings contribute to the growing body of literature on AML and highlight the importance of personalized treatment strategies tailored to the unique genetic and clinical profiles of individual patients. The comprehensive framework we developed for analyzing and interpreting the data proved to be effective in identifying meaningful patterns and associations. This framework, which can be applied to other datasets, offers a systematic approach for the lab to continue exploring the genetic underpinnings of AML and other cancers. By standardizing the analysis process, we aim to facilitate the comparison of results across different studies and datasets, ultimately contributing to a more cohesive understanding of AML and its treatment. As the field of precision medicine continues to evolve, studies like this one will be crucial in guiding the development of more effective, targeted therapies that improve patient outcomes and reduce the mortality associated with this challenging disease. Future research should focus on validating these findings in larger, independent cohorts and on exploring the potential of the identified genes as biomarkers for early detection and treatment stratification. Additionally, further investigation into the molecular mechanisms underlying the observed associations could provide deeper insights into the pathogenesis of AML and inform the design of novel therapeutic strategies aimed at targeting the most critical drivers of disease progression. By advancing our understanding of the factors that influence survival in AML, we hope to contribute to the development of more personalized and effective treatment approaches, ultimately improving the prognosis for patients diagnosed with this devastating disease.

Determining the Role of Codanin-1 in Congenital Dyserythropoietic Anemia Type I

Meira Strauss
Advised under Professor Benny Motro

Congenital dyserythropoietic anemia type I (CDA-1) is a rare type of blood disorder in which erythrocyte production is ineffective. It is an autosomal recessive disease characterized by morphological abnormalities in the chromatin of erythroid precursors. The primary cause of this disorder is mutations in the gene encoding the protein Codanin-1 (CDAN-1). A possible function of CDAN-1 may be to support the integrity of the nuclear envelope. Additionally, CDAN-1 has been shown to act as a scaffold protein and influence the levels and activities of the nuclease C15orf41. The exact function of CDAN-1 both in erythropoiesis and in other tissues is still unknown and is being studied with many different methods and techniques. Analysis of the role of CDAN-1 in cellular functions can help better understand CDA-1 and combat this disease.

The role of CDAN-1 in erythropoiesis was studied using the K562 cell line, a nucleated erythroleukemia cell line derived from a 53 year old female with Chronic myelogenous leukemia. The cells were treated with hemin, a porphyrin which induces differentiation of K562 cells. Cl.3 K562 cells which expressed the CDAN-1-Degron and the control wild type cells were treated with hemin and auxin. The Auxin Inducible Degron system works as follows: an endogenous protein of interest is “tagged” with degron and in the presence of auxin and auxin receptor (OsTIR1) the protein is ubiquitinated and is sent to the proteasome for degradation. Five days after treatment (differentiation time) the cells were stained with CD71 and CD235 antibodies, fixed and underwent FACS. The CD71 antibody is specific for the transferrin

receptor, a receptor which helps immature red blood cells take up iron. Therefore, it is expressed on reticulocytes, and erythroid precursors and indicate the number of cells which have begun to differentiate. In contrast, the CD235 antibody is specific for glycophorin A and glycophorin B, and will bind to more mature erythrocytes. In FACS, each cell passes a laser beam and will fluoresce based on the fluorescent markers it carries. The detector observes and records these signals.

As shown in Figure 1, the results of the FACS experiment may indicate the role of CDAN-1 in differentiation. Tube 2 (left) corresponds to WT cells treated with Hemin and tube 5 (right) corresponds to Cl. 3 cells treated with IAA and Hemin. 42.6% of the WT Hemin cells are found in Q2, the region for positive markers for both CD235 and CD71, while only 30.3% of the Cl.3 IAA Hemin cells are found in this region. This indicates a lower level of differentiation in cells which are presumed to have had lower levels of CDAN-1. Additionally, 28.9% of WT Hemin cells are found in Q4 (the region for only positive markers for CD71- the marker for erythrocyte precursors only) while only 10.4% of the Cl.3 IAA Hemin cells are found in this region. This indicates that there are not only fewer cells which have completed differentiation in CDAN-1 deficient cells but also fewer which have begun differentiation. The experiment was repeated multiple times and the results remain inconclusive of definite correlation, therefore the experiment is still being carried out.

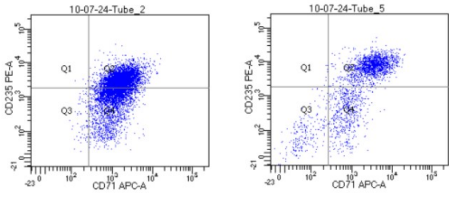


Figure 1: FACS Results

The role of CDAN-1 outside the hematopoietic system was tested using the Cre-Lox P system in mice. Homozygote mice for Coda-floxed were crossed with mice heterozygous for Coda-floxed and heterozygous for SF1-Cre. SF1 is the promoter for Cre and is expressed in the brain, testes, and ovaries. DNA was extracted from 3 week old offspring. PCR was carried out on the DNA samples and the samples were loaded and run in agarose gel.

According to Mendelian Genetics, in this cross 4 genotypes are expected: heterozygous CDAN-1 Floxed, mutant CDAN-1 Floxed, heterozygous CDAN-1 Floxed Heterozygous cre, mutant CDAN-1 Floxed Heterozygous cre. Theoretically, each one of these should make up 25% of the total offspring. However, as shown in Figure 2, the results reflect the existence of other factors; the total absence of homozygotes CDAN1-Floxed SF1-CRE suggest that it is lethal or that Cdan1 and SF1-Cre are adjacent on the same chromosome. The existence, while at lower percentage, of heterozygotes for CDAN1-floxed suggest that the genes are on the same chromosome but are not very close (allowing some recombination), and the absence of any homozygotes CDAN1 -Floxed SF1-CRE confirm that it is lethal.

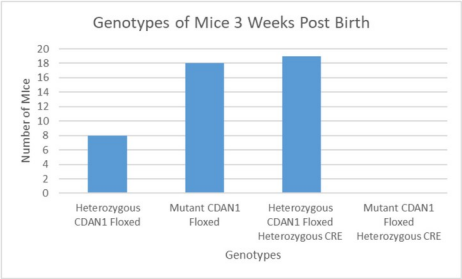


Figure 2: Genotypes of Mice 3 Weeks Post Birth

To examine at which stage Codanin-1 is essential, DNA was extracted from embryos around day 18, PCR was carried out and the samples were run on agarose gel. As shown in figure 3, there were two mice which are mutant

for CDAN-1 flox and heterozygous for SF1-Cre.This indicates that prior to birth, the mice embryos are able to develop without CDAN-1 in certain tissues where Cre is expressed (brain, ovaries, and testes).

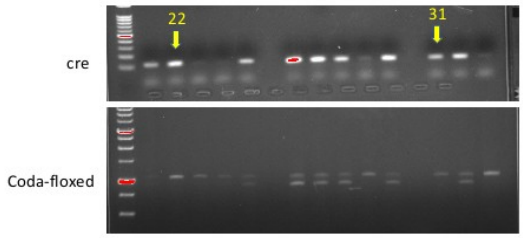


Figure 3: Genotyping of Late Stage Embryos

The FACS experiment, while still ongoing, has shown that CDAN-1 may be essential for erythrocyte differentiation in vitro, and the K562 cell line can serve as a model for CDAN-1/C15 functions. Additionally, using the Cre-Lox P system, it was shown that mice which do not produce CDAN-1 in organs where SF1-CRE is expressed do not survive to adulthood. CDAN-1 may not be essential for embryonic development of the brain, ovaries, and testes.

A Study of Sirt6, the Anti-aging Protein, and its Overexpression in Liver Tissue

Ella Nasser and Alyssa Polonetsky
Advised under Prof. Haim Cohen and MSc students Ron Nagar and Zacharia Schwartz

For the past few centuries, research has successfully focused on understanding and targeting various illnesses and disorders. However, since the natural aging process itself prompts numerous diseases and health concerns, scientists are eagerly searching for methods to extend life and health span. Within the structural makeup of our chromatin lies the enzyme Sirtuin 6, or Sirt6, a member of the 7 mammalian sirtuins. Sirt6 regulates various cellular processes, including DNA repair, metabolism, and inflammation. Most importantly, Sirt6 slows down the cell's aging process through DNA modulation. As an NAD+-dependent enzyme, Sirt6 exhibits deacetylase activity, primarily targeting Histone 3 and removing acetyl groups from specific lysine residues. In this process, the positive charge becomes neutralized and the chromatin closes. Sirt6 is also capable of mono-ADP-ribosyltransferase activity, adding ADP-ribose units to target proteins with the NAD+ substrate. Sirt6 levels decrease as we grow older, and this is likely why various health issues arise. As chromatin becomes more open with age and DNA is over-transcribed, cells begin to lose their identities. Higher Sirt6 levels are necessary to keep the chromatin condensed, closed, and young[1]. New research aims to better understand the mechanisms of Sirt6 and clearly define its specific effects within the cell.

As opposed to previous experimentation in which mice were bred to overexpress Sirt6, the focal point of this project was to work with the hepatocytes of mice that were treated at an older age. To ensure that no other factors interfere with the results, the control group overexpressed green fluorescent protein (GFP) instead. Formalin-fixed paraffin embedding (FFPE) was carried out on the 15 livers that had been harvested in early spring: 9 of Sirt6 overexpression and 6 of GFP. To stain the tissue with a variety of stainings, the samples were cut into 4-micron-wide sections by using a microtome for the preparation of 90 slides (3 sections per slide). Another essential part of the staining process was calibrating various steps. To accomplish this, previously embedded liver samples from 4 groups of mice overexpressing Sirt6 (MOSES mice) were similarly cut and prepared into 24 slides. The groups were young wildtype, old wildtype, young transgenic, and old transgenic.

The first staining calibrated was a baseline H&E (hematoxylin and eosin) stain. Combined with

the Sirius Red stain, a microscope can uncover the potential effect of Sirt6 on fibrosis of the liver. Staining involves dewaxing the slides, rehydrating them by immersing them in various ethanol solutions, and finally incubating them with a stain. The H&E calibration pointed to an ideal immersion time of 2 minutes in hematoxylin and 45 seconds in eosin, but the solvent picric acid for Sirius Red awaits delivery to the lab. Sirius Red stains collagen fibers, and fibroblasts can be identified based on densities within the collagenous matrix. This staining should highlight the collagen levels of the tissue and reveal whether Sirt6 lowers the risk for fibrosis in the liver by directly targeting its fibroblasts.

More results came with the immunofluorescence aspect of this project. Calibrations led to a five-minute staining time for Hoechst, a blue staining of nuclear DNA. Additionally, a one-hour incubation time was established for a mixed blocking and permeabilization solution which contained Triton and donkey serum. Immunofluorescence staining follows the same dewaxing and rehydration protocol but also involves antigen retrieval with tris-EDTA or citrate buffer, use of a hydrophobic PAP pen, dilution of antibodies, and a dark humidity chamber for incubations. The first antibodies used in the immunofluorescence stainings were Sirt6 and Sirt6 EPR to properly detect the enzyme in the treated mice. Successfully staining for Sirt6 would glean insight into the nuclear nature of the protein, and can ensure that it is uniformly distributed within the nuclei and amongst the cells. The efficacies of both antibodies were established via stainings with MOSES mice. Antibody Sirt6 seemed to emit a stronger signal than Sirt6 EPR, but the latter may have bound to the protein more specifically (Figure 1). Both antibodies did not detect any Sirt6 in old wild-type mice, yet this aligns with prior results of our laboratory. Sirt6 appeared to be uniformly distributed throughout the nuclei and in every hepatocyte.

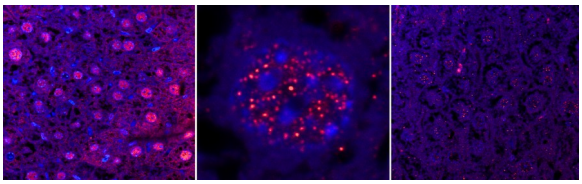


Figure 1: Immunofluorescence stainings of old transgenic liver tissue at 60x magnification. Left to right: Sirt6, Sirt6 (1 nuclei), Sirt6 EPR. DNA was stained with Hoechst (blue) and Sirt 6 was stained with Alexa 647 (red).

The next immunofluorescence stain aimed to assess whether an increase in Sirt6 was directly correlated with a reduction in macrophages and inflammation. Macrophages were weakly detected via antibody F4/80 in two old, wild-type mice. Nevertheless, scientific literature confirms that a weak signal of the F4/80 antibody is indicative of macrophage presence (Figure 2). All of these calibrations provide promising protocols and insights for working with the samples from mice already at an old age.

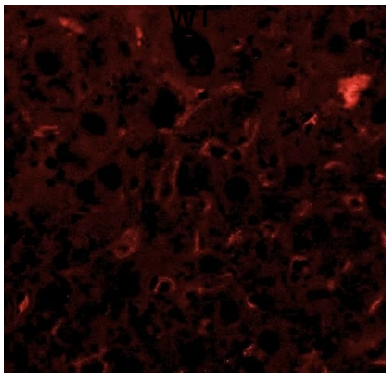


Figure 2: Immunofluorescence staining with F4/80 antibody in liver tissue of old wild-type mice using Cy5 (red) to detect macrophages.

Lastly, a western blot used proteins from liver cell lines of 8 mice overexpressing Sirt6 and 7 mice overexpressing GFP to determine the involvement of Sirt6 in different metabolic pathways. After the gel was run, the proteins were transferred to a nitrocellulose membrane and then treated with Ponceau stain and TBST

washes. The membranes were incubated with 5 different primary antibodies, including AMPK, pAMPK, S6, pS6, and C-JUN, followed by secondary antibodies. Chemiluminescence signals were developed and captured using ECL substrate and LAS4000. mTOR and AMPK reflect two parallel pathways involved in proliferation and conservation, respectively. If Sirt6 successfully increases AMPK, there would be higher levels of pAMPK. On the other hand, if Sirt6 decreases the mTOR pathway, there would be lower levels of pS6. As a control, total protein levels were measured and ratios were derived. Results showed that Sirt6 overexpression did not increase the activity of the AMPK pathway nor decrease that of the mTOR pathway when compared to GFP-producing mice (Figure 3). Additionally, levels of pioneer transcription factor c-Jun also did not yield a significant discrepancy between Sirt6-overexpression and GFP-producing mice. This indicates that the role of Sirt6 in condensing chromatin structure seems not to involve c-Jun protein levels, but further blotting of phospho-c-Jun should provide more clarity.

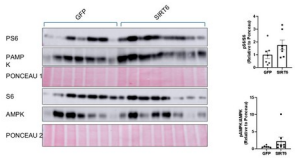


Figure 3: Western blot and quantitative analysis of phosphorylated and total AMPK and S6 protein levels in mice treated at old age expressing Sirt6 (n=8) and GFP (n=7). No significant differences were observed; outliers were excluded.

With 90 slides of samples carefully prepared as well as promising results until this point, the race to understand the science behind longevity and Sirt6 is well on its way. Extensive experimentation with more staining and blotting of Sirt6 in the treated samples will only further the discoveries necessary to complete the ultimate picture.

[1] Dong X. Charlie, Cells, 12(4), 663 (2023).

PHYSICS, CHEMISTRY, AND MATHEMATICS



Daniel Lipshitz, Jonathan Beer, Eliana Fishman, Rachael Farkas, Sophia Purow, Elana Rosenblatt

The Varieties of n-Dimensional Lie Algebras and Their Components

Jonathan Beer
Advised under Dr. Eyal Subag

Introduction

A Lie Algebra is a vector space equipped with a skew-symmetric bilinear map called a Lie Bracket which satisfies the Jacobi Identity [X,[Y,Z]] + [Z,[X,Y]] + [Y,[Z,X]] = 0. Lie Algebras are intimately connected to Lie Groups (Groups which are also manifolds); the category of simply connected Lie Groups is identical to that of Lie Algebras, and there is a sense in which Lie Algebra maps are the derivatives of maps from simply connected Lie Groups to Lie Groups, called the Lie Homomorphism Theorem. The functor that takes a Lie Group to its corresponding Lie Algebra is called Lie, and is among the fundamental tools for studying Lie Groups.

The Variety of Lie Algebras

Given an ordered basis on an n-dimensional vector space V, the space of bilinear maps on V (isomorphic to $V \otimes V^* \otimes V^*$) can be given a basis which allows the bilinear maps to be expressed as an n^3 dimensional tensor. These are called

the *structure constants* of the map. In this setting, the Jacobi Condition amounts to a set of quadratic conditions on the coordinates of a bilinear map, and the Lie Algebras thus form an algebraic subset of this n^3 dimensional space called the *Variety of Lie Algebras*. A paper by Kirillov and Neretin[1] investigated the low dimensional (dimension 3 and 4 in particular) varieties over the complex numbers. We hoped to extend several of their findings to higher dimensions and other algebraically closed fields. Moreover, the quotient of this space by the action of GL(V) (given in Eq. 1) gives a variety of non-isomorphic Lie Algebras. Having an injection or bijection into the the variety makes studying its properties much easier.

$$\begin{bmatrix} \tilde{c}_{12}^2 \\ \tilde{c}_{13}^3 \\ \tilde{c}_{12}^1 \\ \tilde{c}_{23}^3 \\ \tilde{c}_{13}^1 \\ \tilde{c}_{23}^2 \\ \tilde{c}_{23}^1 \\ \tilde{c}_{13}^2 \\ \tilde{c}_{12}^3 \end{bmatrix} = \begin{bmatrix} N_{22}R_{33} & N_{23}R_{23} & N_{21}R_{33} & N_{23}R_{13} & N_{21}R_{23} & N_{22}R_{13} & N_{21}R_{13} & N_{22}R_{23} & N_{23}R_{33} \\ N_{32}R_{32} & N_{33}R_{22} & N_{31}R_{32} & N_{33}R_{12} & N_{31}R_{22} & N_{32}R_{12} & N_{31}R_{12} & N_{32}R_{22} & N_{33}R_{32} \\ N_{12}R_{33} & N_{13}R_{23} & N_{11}R_{33} & N_{13}R_{13} & N_{11}R_{23} & N_{12}R_{13} & N_{11}R_{13} & N_{12}R_{23} & N_{13}R_{33} \\ N_{32}R_{31} & N_{33}R_{21} & N_{31}R_{31} & N_{33}R_{11} & N_{31}R_{21} & N_{32}R_{11} & N_{31}R_{11} & N_{32}R_{21} & N_{33}R_{31} \\ N_{12}R_{32} & N_{13}R_{22} & N_{11}R_{32} & N_{13}R_{12} & N_{11}R_{22} & N_{12}R_{12} & N_{11}R_{12} & N_{12}R_{22} & N_{13}R_{32} \\ N_{22}R_{31} & N_{23}R_{21} & N_{21}R_{31} & N_{23}R_{11} & N_{21}R_{21} & N_{22}R_{11} & N_{21}R_{11} & N_{22}R_{21} & N_{23}R_{31} \\ N_{12}R_{31} & N_{13}R_{21} & N_{11}R_{31} & N_{13}R_{11} & N_{11}R_{21} & N_{12}R_{11} & N_{11}R_{11} & N_{12}R_{21} & N_{13}R_{31} \\ N_{22}R_{32} & N_{23}R_{22} & N_{21}R_{32} & N_{23}R_{12} & N_{21}R_{22} & N_{22}R_{12} & N_{21}R_{12} & N_{22}R_{22} & N_{23}R_{32} \\ N_{32}R_{33} & N_{33}R_{23} & N_{31}R_{33} & N_{33}R_{13} & N_{31}R_{23} & N_{32}R_{13} & N_{31}R_{13} & N_{32}R_{23} & N_{33}R_{33} \end{bmatrix} \begin{bmatrix} c_{12}^2 \\ c_{13}^3 \\ c_{12}^1 \\ c_{23}^3 \\ c_{13}^1 \\ c_{23}^2 \\ c_{23}^1 \\ c_{13}^2 \\ c_{12}^3 \end{bmatrix}$$

Equation 1: The action of $GL(C^3)$ on the structure constants of the variety corresponding to C^3

The Symmetric Component

The paper by Kirilov and Neretin[1] used a particular map from Sym(V)⊕V to Hom(Λ^2V,V) (which in particular is an isomorphism in 3 dimensions, given in Eq. 2) to study the 3-dimensional Variety of Lie Algebras. Unfortunately, this map cannot possibly generalize to an isomorphism in arbitrary dimensions due to dimension mismatch.

$$[\cdot,\cdot]_T = T([T^{-1}(\cdot), T^{-1}(\cdot)])$$

Equation 2: The action of T \in GL(V) on a bilinear map [X,Y]

We investigated an obvious way to extend this map. We obtained that the obvious extension gives an injection into the variety of Lie Algebras in dimension 4, and not in dimension 5 and above.

$$\begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ s^{12} \\ s^{13} \\ s^{23} \\ s^{11} \\ s^{22} \\ s^{33} \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & -2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2 \end{bmatrix} \begin{bmatrix} c_{12}^2 \\ c_{13}^3 \\ c_{12}^1 \\ c_{23}^3 \\ c_{13}^1 \\ c_{23}^2 \\ c_{23}^1 \\ c_{13}^2 \\ c_{12}^3 \end{bmatrix}$$

Equation 2: The bijective map between Sym(V)⊕V and Hom(Λ2V,V)

Conclusions and Future Research

The symmetric component given in 3 and 4 dimensions lines up somewhat with the Lie Algebras given by the orthogonal algebra relative to some symmetric bilinear form. This map obviously generalizes to an injection into the variety of Lie Algebras of arbitrary dimension, but it is less useful for investigating the larger variety of Lie Algebras. More methods for investigating components of the variety of Lie Algebras which work in arbitrary dimensions would be useful for studying this variety in higher dimensions, particularly higher than 4 where the varieties grow increasingly poorly understood.

[1] Kirillov, Alexandre & Neretin, Yuri. (1987). The Variety An of n-dimensional Lie Algebra Structures. Am. Math. Soc. Transl. 137. 10.1090/trans2/137/03.

Discrete Curve Shortening Flows and Geometric Properties through Fourier Decomposition of Planar Polygons

Daniel Lipshitz
Advised under Dr. Menachem Lazar

Our study investigated the discrete analog of Curve Shortening Flow (CSF) for polygons, focusing on the use of the discrete Fourier transform to analyze and evolve the geometric properties of discrete curves. CSF, a process in differential geometry, is known for evolving

smooth curves towards circular shapes by moving points inward along their normals at a rate proportional to their curvature at each point. While previous work has shown CSF shrinks embedded, non-overlapping, curves into convex round points, adapting this flow for discrete curves like polygons presents challenges due to the lack of continuous curvature.

We investigated possible flows by focusing on the Fourier transforms of polygons. The Discrete Fourier Transform (DFT) decomposes each polygon into a unique sum of n regular polygons, each with a certain magnitude and phase. These regular polygons, corresponding to columns of the DFT matrix, represent traversals of nth roots of unity, or n evenly spaced points on the unit circle, at a kth periodic frequency. By defining a region in the this domain, we aim to establish a discrete flow that drives polygons towards convex, round points.

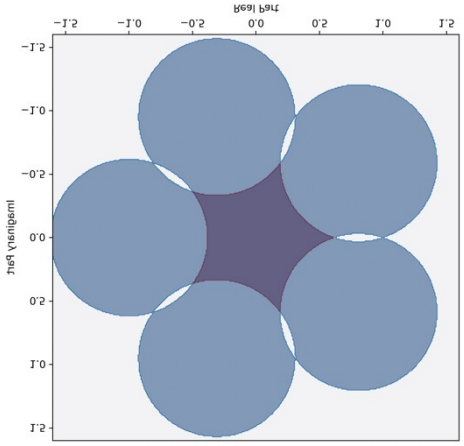


Figure 1: Convex and simple region found as slice of higher dimensional region by fixing all but one coefficient

In the frequency domain we were able to express desirable geometric properties through simplified equations, as well as describe shapes based on their underlying geometric properties independent of symmetries which many share. We examined existing flows, such as the midpoint flow, which evolves polygons towards regular n-gons by adjusting vertex positions based on neighboring edges. While effective, this flow does not achieve the finite-time shrinking characteristic of continuous CSF. To address this, we used a discrete flow that better approximates CSF by normalizing vertex movements and ensuring that the polygon's area decreases at a constant rate. We began to prove how this flow is maintains convexity and simplicity throughout the evolution, thereby bringing all embedded polygons to round convex shapes in finite time.

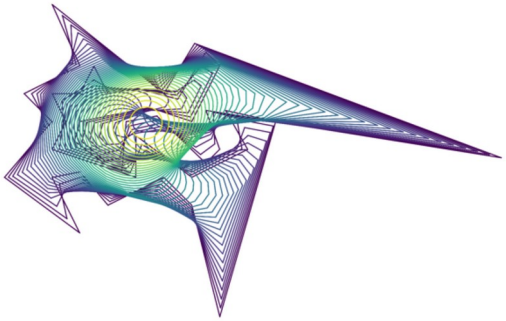


Figure 2: Selected time steps under proposed flow bringing curve to round point

Finally, we work analyzed singularities in the flow, particularly when polygons experience cusp-like behavior or when edges become parallel or anti-parallel. The proposed flow is shown to navigate these singularities, ensuring continuous evolution except in the case of figure-eight shapes, which halt the flow. Through these investigations, we aim to provide a robust framework for the discrete evolution of polygons, preserving the geometric properties that characterize continuous CSF.

[1] B. Chow and D. Glickenstein, The American Mathematical Monthly, **114(4),**316–328, (2007).

Machine Learning-Guided Discovery of Multinary Metal Oxides for Sustainable Energy Applications

Rachael Farkas
Advised under Dr. Hannah-Noa Barad

The drive for sustainable energy has led to a growing interest in discovering advanced materials that can support eco-friendly technologies. The lab under Dr. Barad at Bar-Ilan University focuses on developing multinary metal oxides—materials composed of multiple elements with unique properties suited for catalytic applications. These materials are crucial in OER- Oxygen Evolution Reaction which are helpful to producing clean fuels. By employing machine learning and computational methods, the lab seeks to enhance the efficiency of material discovery and optimization, using data-driven approaches to navigate the complex landscape of potential material combinations and properties.

My objective was to analyze properties of metal oxides, such as overpotential and band gaps, in order to identify correlations that can be used to train machine learning models for predicting material performance and optimizing properties.

The research methodology combined academic literature review, specialized training, and practical application of machine learning techniques to explore multinary metal oxides for sustainable energy applications.

Literature Review: I reviewed academic papers on multinary metal oxides, focusing on catalysts for CO2 reduction and water splitting.

Machine Learning for Materials Research Bootcamp: To enhance my computational skills and address my limited chemistry background, I attended a bootcamp covering machine learning techniques in materials science, such as feature selection, model training, and validation.

Materials Project: I used the Materials Project database to study chemical and structural properties, helping identify promising materials based on computed properties.

Data Analysis and Visualization: I analyzed data with machine learning models, creating visualizations to explore relationships between material properties and catalytic activities, identifying trends and correlations.

This integrated approach enabled me to contribute effectively to the research at Bar Ilan University, bridging knowledge gaps and advancing the discovery of new materials.

Initial analysis has revealed promising correlations between metal oxide properties. Notably, certain environmental conditions seem to correlate with increased material overactivity, possibly tied to band gap characteristics. We also found trends indicating that crystal structure might influence electronic properties. Furthermore, while these findings are encouraging, additional analysis is required to confirm and refine these insights. Future work will aim to better predict material performance and identify oxides with lower overpotential for sustainable energy applications.

The analysis involved examining relationships between metal oxide properties using data from both experimental and computational sources. I employed statistical techniques to explore correlations among key characteristics such as band gaps, catalytic activity, and crystal structure. Data visualization tools, including graphs and plots, were used to identify trends and patterns in these properties' interactions.

Using Python, I implemented code to color-coordinate the data points based on crystal structure or environmental factors, allowing for

the visualization of clusters and highlighting distinct groupings. This organized approach not only facilitated the identification of initial correlations but also provided a framework for future analyses. By visually distinguishing these clusters, we can more easily observe and explore patterns, enabling a more nuanced understanding of how different properties and structures interact.

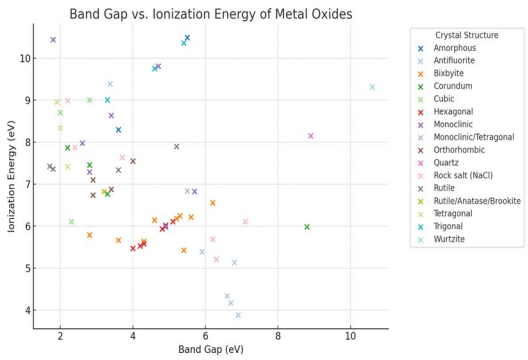


Figure 1: The starting trends between Band Gap and Ionization Energy of Metal Oxides, based on the Crystal Structure

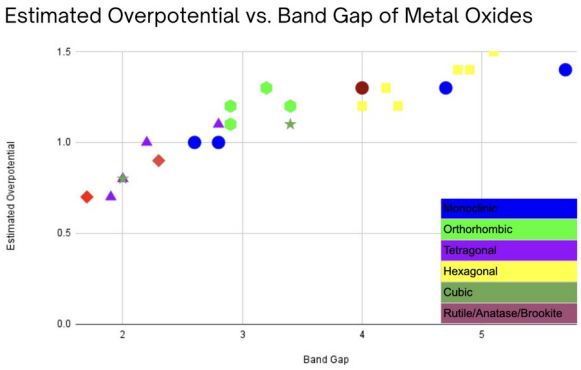


Figure 2: Starting trends comparing Estimated Overpotential against the Band Gap of Metal Oxides characterized by their Crystal Structure Enviornment.

Over the past month, this research has advanced our understanding of multinary metal oxides and their properties. By integrating machine learning with data analysis, we've started to uncover important correlations between material characteristics. These preliminary findings provide a valuable basis for deeper investigation. Beyond the technical insights, this project has been an enriching learning experience, enhancing my expertise in applying computational techniques to materials science and developing predictive models. The knowledge gained will be instrumental in refining our approach and driving future research efforts toward more effective and sustainable energy solutions.

Inhibition of P53 aggregation with cyclic D,L-α-peptide

Eliana Fishman
Advised under Professor Shai Rahimipour and PhD student Roshan Roy

Skin-derived precursor cells (SKPs) are neural crest stem cells that are found in adult epidermis cells[1]. Notably, SKPs overexpress Parkinson's disease-associated α-synuclein (α-syn) just as neurons do, which can potentially realize the higher-than-average risk of melanoma cases in Parkinson's disease (PD) patients. In addition, in both melanoma and PD patients, there is also an overexpression of P53 in the cells[2].

P53 is a tumor suppressor gene which controls cell apoptosis. Mutations in P53 lead to overexpression which results in aggregation of the protein and loss of its function. Mutations in P53 have been found in over 50% of human cancers. Previous research has indicated a strong link between PD and melanoma diseases1. In PD, α-syn accumulates and aggregates in the substantia nigra of the brain and causes premature neuronal cell death. Similarly, aggregation of P53 in the melanocytes of melanoma patients leads to loss of function, stopping the apoptosis process and causing uncontrolled cell proliferation. Previous studies have found that cyclic D,L-α-peptides (CP-2) can interact with early forms of many amyloidogenic protein including α-syn, amyloid β, and tau to inhibit their aggregation and toxicity[4]. Therefore, current research focused

on assessing whether cyclic D,L-α-peptides can also interact with P53 and inhibit its aggregation to potentially restore the function of the tumor suppressor protein and increase tumor cell apoptosis.

Two aggregating fragments of P53 with sequence of ILTIITLE and RRPILTITLEDSSGNLLGRNSFEVR were prepared using the solid phase peptide synthesis method. A chlorotrityl resin was used as the solid support and the synthesis was based on the Fmoc protocol. HBTU was used as the coupling reaction and DIPEA was used as the base. The synthesis was repeated twice, once with 5-folds and the other with 10-folds excess of the amino acids and the coupling reagent. The peptides were cleaved from the resin using trifluoroacetic acid (TFA) solution and the purity of the synthesized peptides was determined using high-performance liquid chromatography (HPLC) and mass spectrophotometry (Figures 1, 2)

The analytical HPLC (Figure 1) shows that preparative purification of the peptide was successful and the peptide is pure (>95%).

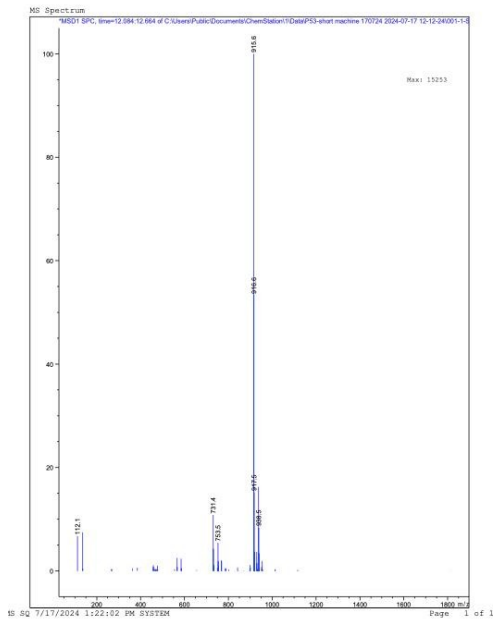


Figure 2: Mass Spectrophotometry of crude synthesized ILTTIITLE

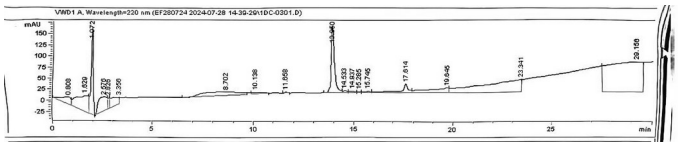


Figure 1: HPLC Profile of pruified ILTIITLE

The mass spectrophotometry results suggest that mass of the peptide is 914.57 g/mol , which matches the theoretical mass of ILTIITLE.

We found that very high concentrations of acetonitrile and prolonged sonication time were required to dissolve the peptides. This indicates the strong hydrophobic character of the peptides and potentially their self-assembly capability. The aggregation of the peptides in aqueous solution was also noticed in HPLC chromatograms as broad peaks. The high aggregation propensity of the P53-derived sequence in solution demonstrates why it is so frequently found aggregated in the cells of cancer patients. The short P53 was successfully purified by semi-preparative HPLC to yield the pure peptide. The pure peptide typically eluted from the HPLC column (C18) at 60% of acetonitrile in H2O.

Once purified, the shorter sequence of P53 was tested by thioflavin T (ThT) assay to determine the lowest concentration of peptide that

aggregates in aqueous solution. The interaction of amyloid fibrils with ThT causes an increase in fluorescent signal that can be detected by a fluorescent plate reader. The preliminary ThT results (Figure 3) suggested that the short P53 aggregated at the concentration of 0.175 mg/ml, however these results should be repeated.

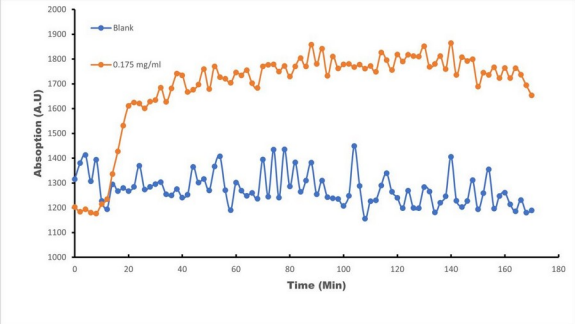


Figure 3: Preliminary ThT results for aggregation of ILTIITLE at 0.175 mg/ml

Future studies will include optimization of the ThT aggregation protocol to discover the lowest concentration of P53 that aggregates. With this knowledge, the next steps would be to test the inhibitory effect of CP-2 on aggregation of P53 analog using the optimize ThT assay. These experiments could also be repeated using the longer sequence of P53 that had previously been synthesized. Next, cell culture experiments will be performed using different cancer cell lines to determine the ability of CP-2 to inhibit P53 aggregation in vitro and consequently increase the apoptosis of the cancer cells.

[1] Bose A, Petsko GA, Eliezer D. Parkinson's Disease and Melanoma: Co-Occurrence and Mechanisms. J Parkinsons Dis. 2018;8(3):385-398. doi: 10.3233/JPD-171263. PMID: 29991141; PMCID: PMC6130416.

[2] Bataille A, Leschiera R, L'Hérondelle K, Pennec JP, Le Goux N, Mignen O, Sakka M, Plée-Gautier E, Brun C, Oddos T, Carré JL, Misery L, Lebonvallet N. In vitro differentiation of human skin-derived cells into functional sensory neurons-like. Cells. 2020;9, 1000. PMID: 32316463

[3] Luo Q, Sun W, Wang YF, Li J, Li DW. Association of p53 with neurodegeneration in Parkinson's Disease. Parkinsons Dis. 2022, 2022, 6600944, PMID: 35601652

[4] Richman M., Wilk S., Chemerovski M., Wärmländer S. K., Wahlström A., Gräslund A., Rahimipour S. In vitro and mechanistic Studies of an anti-Amyloidogenic self-assembled cyclic D,L-alpha-peptide architecture. J. Am. Chem. Soc. 2013, 135, 3474, PMID: 23360549

Platinum Nanoparticles Catalyst for Oxygen Reduction Reaction in Hydrogen Fuel Cells

Elana Rosenblatt
Advised under Prof. Lior Elbaz and MSc student Eliana Lebowitz

As global energy consumption increases, the pollution and CO_2 emissions that accompany the burning of fossil fuels increase as well. In order to meet energy demands while protecting the environment, scientists are working to develop sources of zero-emissions energy. One such source of energy is hydrogen fuel cells. In a hydrogen fuel cell, chemical energy is converted directly into electrical energy through a reduction – oxidation (redox) reaction where a fuel (hydrogen) is oxidized, and oxygen is reduced. As the only product of this reaction is water, fuel cells are a zero-emission energy source.

The power generated by a fuel cell is dependent both on the voltage as well as the current generated. Due to the slow kinetics of the oxygen reduction reaction (ORR), voltage is lost and fuel cells are not able to generate as much power as is theoretically possible. In order to maximize the efficiency of fuel cells, electrocatalysts are being developed for the ORR to speed up the reaction and reduce the energy loss. As common catalysts that are available for this reaction are expensive, new catalysts are being developed and studied in order to allow for the commercialization of fuel cells. In this experiment, a Pt nanoparticle catalyst was synthesized in order to test its catalytic abilities in the ORR reaction. The Pt nanoparticles were synthesized on a carbon support through wetness incipient

impregnation and reduction using H_2 gas. Chloroplatinic acid and XC72 (carbon black) were dissolved in acetone and the resulting solution was allowed to evaporate on a hot plate before being reduced using H_2 gas in an oven at 80 ℃. The resulting catalyst was characterized using X-Ray Diffraction (XRD), High Resolution Scanning Electron Microscopy (HRSEM) and Electron Dispersive X-Ray Analysis (EDAX). Its catalytic abilities were studied using cyclic voltammetry (CV) and rotating disk electrode (RDE) voltammetry in a half cell. Following these procedures, a full fuel cell was constructed using a commercial catalyst in order to demonstrate the properties of a fuel cell.

XRD: The XRD graph, depicted in Figure 1, shows peaks at 39.1°, 46.1° and 67.2°. Based on the literature, these peaks correspond to the (111), (200), and (220) planes in the crystal structure, respectively, which are characteristic of a face centered cubic crystal. As platinum crystalizes in a face centered cubic crystal, this indicates the presence of Pt in the sample. The XRD data was also used to determine the mean Pt particle size, based on the peak width and the Scherrer equation. The average particle size was calculated to be 2.76 nm. This particle size was consistent with the range of particle sizes present in the data.[1]

HRSEM and EDAX: The SEM was not high enough resolution to depict the Pt nanoparticles, and therefore only depicted the carbon support. EDAX provided a chemical mapping of the sample, qualitatively depicting the dispersion of the particles. The EDAX scan indicates that the Pt was dispersed throughout the sample.

CV: Figure 2 depicts the cyclic voltammogram (CV) of the half cell. Figure 2 is labeled with the various reactions that begin at each voltage differential in the half cell. The ECSA (electrochemical surface area) of the electrode, normalized by the mass of Pt on the electrode, was calculated to be 10.3 m^2/g, based on the average area of the adsorption and desorption peaks, shown shaded in the figure. This is greater than the geometrical surface area, as expected. The larger the surface area, the more effective the catalyst.

RDE: Figure 3 depicts the plot from the rotating disk electrode voltammogram. Unlike in CV, in RDE voltammetry, diffusion is controlled, and the current is limited by the rate of diffusion. Therefore, the limiting current increases as the speed of rotation increases. Using the limiting current and the Levich equation, the number of electrons transferred in this redox reaction was calculated to be an average of 4.58 electrons, which is close to the desired number of 4.00 electrons. A 4.00 electron reaction is desired as it allows for maximum current, and therefore maximum power, to be generated.

Fuel Cell: Figure 4 depicts the IV polarization curve of a fuel cell constructed with a commercial catalyst. As indicated in Figure 4, there are 3 regions in the IV curve. Due to mixed potential, the open circuit voltage of the fuel cell is lower than the theoretical voltage of the fuel cell, determined by reduction potentials. The first region is the activation region, in which voltage is lost due to the kinetics of the reaction, and the activity of a catalyst can be assessed based on the slope of the curve in this region. The second region is the ohmic region, in which the current and voltage are only limited by resistance, following Ohm's law. In the third region, the mass transport region, the current is limited by the speed of diffusion.

The XRD scans confirmed the presence of Pt in the sample, and indicated that the average particle size was 2.76 nm. The EDAX chemical mapping indicated that the Pt was evenly dispersed throughout the sample. The electrochemical surface area was found to be

10.3 m^2/g of Pt, which is larger than the geometrical surface area, as expected. Based on the RDE of the half cell and the Levich equation, it was confirmed that four electrons were transferred in the reaction, which maximizes the current and allows for maximum power. The IV polarization curve of the fuel cell made using a commercial catalyst demonstrated the polarization in a fuel cell. These results indicate that Pt nanoparticles were successfully synthesized, and that they can be used as an effective electrocatalyst in the ORR.

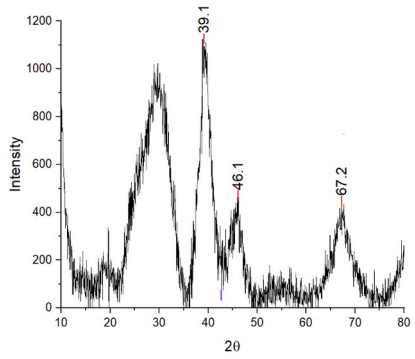


Figure 1: XRD Graph of Synthesized PtC Catalyst

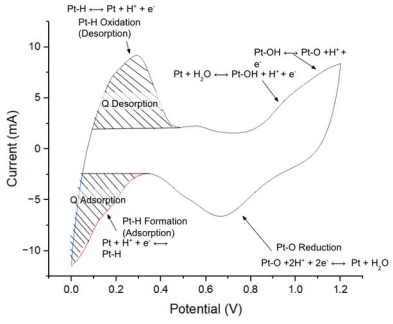


Figure 2: Cyclic Voltammogram in Half Cell Using Synthesized Catalyst[2]

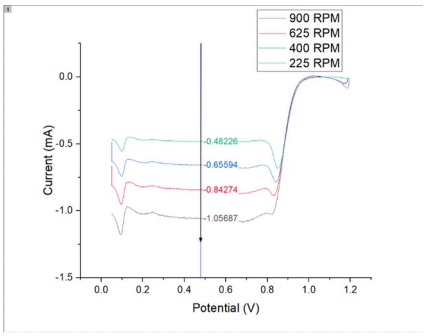


Figure 3: RDE Voltammogram in Half Cell Using Synthesized Catalyst

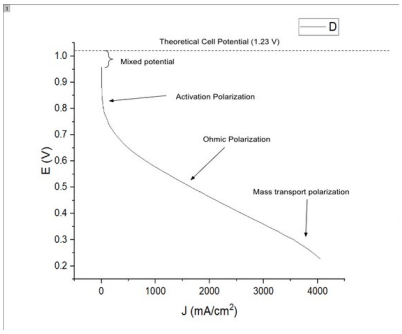


Figure 4: IV Polarization Graph of a Commercial Catalyst

[1] M.A. Shah, Scientia Iranica, **19**, 964-966, (2012).
[2] P. Wongbua-ngam, et al. International Journal of Hydrogen Energy, **44**, 12108-12117, (2019)

Acoustic Analogues of Black Holes and Gravitational Waves

Sophia Purow
Advised under Prof. Patrick Sebbah and Master's Student Jiaoqing Wang

Gravitational waves are ripples in spacetime caused by cosmic events like merging black holes. These waves are often linked to the photon sphere, a region around a black hole where photons can be trapped in orbit by strong gravitational forces. After a merger, the black hole vibrates in specific patterns known as quasinormal modes, influenced by the photon sphere. These modes provide insights into the black hole's mass, spin, and structure. However, studying these vibrations directly is challenging due to the massive scale and distance of black holes.

To overcome observational limitations, we use analogues to mimic the behavior of gravitational waves. The study aims to observe acoustic analogues by vibrating metallic curved shells to attempt to replicate the behavior of waves near a black hole and "listen" to the waves they emit. We want to understand wave dynamics through a controlled environment and to explore whether flexural waves, bending vibrations that occur in structures such as plates or shells, are a valid analogue of the this. We specifically studied the Reissner-Nordström Black Hole (RNBH) model, a charged, non-rotating black hole, and the Schwarzschild Black Hole, a non-charged, non-rotating black hole.

The experimental setup involved 3D printing a metallic elastic shell with a carefully designed curved surface, designed to mimic the spacetime curvature around a black hole, and was created using selective laser melting technology. We propagated an input acoustic signal, typically a short pulse or a continuous wave, through the curved surface, which led to similar wave behaviors that are observed near black holes. The spatiotemporal profile of the velocity field generated by the propagating acoustic waves was recorded using laser vibrometers, which allowed for precise measurement of vibrations across the surface. The vibrations formed from these waves led to the observation of specific modes, such as photon sphere modes (PS modes). We found that the vibration modes are tightly confined in a circular pattern near the photon sphere, with minimal energy loss in that region.

COMSOL Multiphysics software was used to replicate the physical experiment with the vibration of carefully designed models (Figure 1), which replicated the curved surface from the experiment. Eigenfrequency studies were conducted, calculating resonant frequencies and analyzing displacement distributions across various eigenmodes. To simulate realistic wave behavior, Perfectly Matched Layers (PML) were added at the boundaries to reduce reflections at the boundaries and even eliminate them. This eliminates cavity modes resulting from the reflections at the edges of the sample.

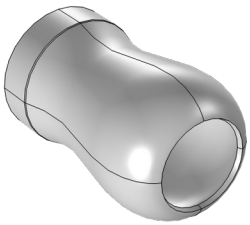


Figure 1: RNBH Simulated Design

Through the simulation, we were able to observe and characterize the PS modes, assess their Q-factors (quality-factors), and explore the effects of boundary absorption on wave confinement.

The Q-factors, representing the ratio of stored energy to energy loss, for each eigenfrequency were calculated, where a higher Q-factor indicated lower energy loss. We aimed to identify modes with high Q-factors, which are indicative of stable wave behaviors in the context of the photon sphere. Photon spheres have minimal energy loss, which would mean Q-factors approaching infinity. By focusing on eigenmodes with near-infinite Q-factors, we identified the PS modes in the model. The simulation provided images and exact coordinates for the photon spheres, revealing stable modes with a single ring, as shown in Figure 2. Repeating the process, allowed us to map the locations of the photon spheres. We found that the eigenfrequencies were evenly spaced apart for PS modes, with the mode's location increasing as the eigenfrequency increased. A graph (Figure 3) was produced to represent the relationship between the locations of the PS modes and their corresponding eigenfrequencies.

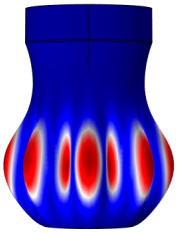


Figure 2: PS Mode of Simulated RNBH

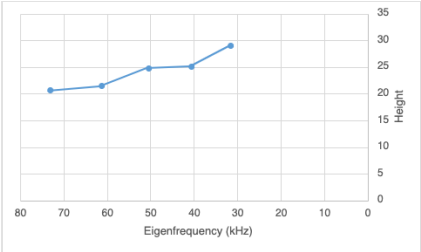


Figure 3: Graph of Eigenfrequency vs Height of PS Modes of SImulated RNBH

During the physical experiments, by creating videos of the vibrations, we were able to see the modes at various eigenfrequencies, which also gave us the locations of the photon spheres with their corresponding eigenfrequencies. The stable PS modes usually had one ring, as shown in Figure 3, supporting the notion that the modes were able to be tightly confined. When testing the Schwarzschild model, we performed experiments with and without an absorber. The absorber aimed to replicate the PML condition from the simulation by reducing reflections at the boundary. With the absorber, fewer modes were detected, supporting the notion that the boundary simulated an open infinite system. Although the absorbers increased energy dissipation, it also allowed for the selective survival of certain modes, particularly the PS modes, which were best confined in the presence of absorbers. This revealed that boundary conditions play a critical role in determining the overall energy dissipation, emphasizing the importance of these conditions in acoustic device design.

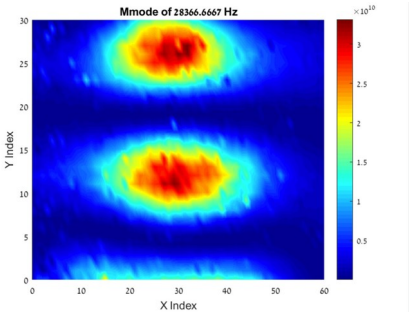


Figure 4: PS Mode at 28366.7 Hz for Long RNBH

In conclusion, this study supports the use of acoustic analogues to explore wave behavior and vibration modes near black holes. The results demonstrate that waves can be tightly confined within a region analogous to the photon sphere and High-Q modes were identified, showing minimal energy loss in acoustic systems. The experimental findings were consistent with the simulations, validating the theoretical models. By vibrating metallic curved shells, we observed flexural waves that may serve as valid analogues for gravitational waves, particularly in the context of the Reissner-Nordström and Schwarzschild black hole models. We successfully identified and characterized vibration modes that exhibit behaviors analogous to those observed near black holes. The inclusion of absorbers highlighted the effect of wave reflections on mode stability, allowing for more accurate observations of wave behaviors. This underscores the significant role of boundary conditions in acoustic device design and the dynamics of wave propagation. Further research is needed to further validate these analogues and expand the exploration of wave dynamics in more complex black hole metrics. Overall, this research bridges theoretical models with practical applications, enhancing our understanding of wave behaviors in curved spaces and offering new opportunities to explore gravitational phenomena in controlled laboratory settings.