UCLA Samueli School of Engineering

# 2023 SURP RESEARCH JOURNAL

*Summer Undergraduate Research Program*

# TABLE OF CONTENTS

Student Research Abstracts and Posters



## BRUCE DUNN

Interim Dean

# INTERIM DEAN'S MESSAGE

The UCLA Samueli School of Engineering's Summer Undergraduate Research Program (SURP) provides a real-world research experience in a wide range of engineering and physical science fields. Through this program, undergraduate students hailing from diverse backgrounds and experiences have an opportunity to conduct research in our faculty's laboratories under their supervision.

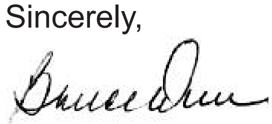
Students in this year's program have been able to:

- Conduct on-campus research in a cutting-edge field at a world-renowned research institution
- Meet and network with a community of peers who share similar goals and interests
- Create a professional scientific poster and publish a research abstract
- Learn to communicate research outcomes and present a detailed project summary
- Gain a competitive experience for applying to graduate schools
- Learn how to impact their communities as engineers and scientists

Fifty-six undergraduate students were selected to join the 2023 SURP cohort and mentored by 37 faculty members across five academic departments. UCLA Samueli is committed to fostering a more equitable, diverse and inclusive community. More than 28% of this year's SURP participants are women, 4% are from underrepresented populations and 7% are first-generation students.

Creating new knowledge is a challenging but meaningful endeavor, and these high-performing students have done an outstanding job working through the rigors of scholarly research. They should be very proud of all that they have accomplished in a short time this summer. I encourage you to peruse our brochure and learn about their research.

Sincerely,



Bruce Dunn


**FACULTY ADVISOR**

**Chee Wei Wong**

**DAILY LAB SUPERVISOR**

**Jaime Flor**

**DEPARTMENT**

**Electrical and Computer Engineering**

**ADRIAN ROZARIO**

Electrical and Computer Engineering
1st Year, UCLA


# Optomechanics in Precise Motion Sensing Applications

## ABSTRACT


Cavity optomechanics is a rapidly expanding field offering innovative approaches for numerous technologies utilizing sub-wavelength light-matter interactions. Optomechanical detection is based on electromagnetic phenomena and ultralow-noise laser readout, rather than traditional techniques such as capacitive or piezoelectric sensing which have higher Johnson electrical noise, consequently achieving sizable higher sensitivities, even at the quantum backaction limits. This high-sensitivity and precision is pertinent in near-term frontier inertial measurement units, which use gyroscopes and accelerometers to measure position and velocity absent external signals acting as references. Inertial measurement is vital in applications from vehicle navigation to motion sensing in smartphones and wearable devices. The operation of the designed optomechanical accelerometer relies on a slotted photonic crystal, which localizes light at specific frequencies determined by its photonic bandgap.

Under acceleration, the slot dimensions shift, causing a corresponding change in the resonant frequency band of trapped light. These shifts show up as measurable changes in the physical oscillation frequency. In this project, I focus on integrating the components of the inertial measurement unit, by modeling suitable housing to be used for navigation in the field. This work compared our optomechanical accelerometer measurement data against leading-edge electronic counterparts. With considerations critical to performance, such as spacing, optical fiber positioning, and heat-weight distributions of required components, I have designed an arrangement with the necessary power supplies, circuits and measurement devices for a near-term flight mission. This assembly transitions from laboratory-scale testing to assessments in the field, effectively increasing the technological readiness level of the system being developed.

# Cavity Optomechanics in Precise Motion Sensing Applications

UCLA Samueli School of Engineering
SUMMER UNDERGRADUATE RESEARCH PROGRAM

Adrian Rozario, Dr. Jaime Flor, Prof. Chee Wei Wong
Mesoscopic Optics and Quantum Electronics Laboratory
Department of Electrical and Computer Engineering - University of California, Los Angeles

## Introduction

- **Inertial measurement unit** (IMU): Measures forces, orientation shifts with gyroscopes, accelerometers
- Measurements may be integrated to calculate velocity, position without external signals or references.
- Navigational capabilities are critical for consumer, industry-grade, military applications *E.g. Smartphones, aircraft, satellites*

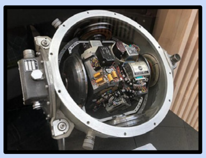


*Figure 1: IMU in the Apollo spaceship. Gimbals are mounted towards the outside, accelerometers towards the center.*

## Cavity Optomechanics

- **Optomechanics**: Development of devices with light-matter interactions at sub-wavelength length scales
- **Slotted photonic crystal** is used to guide laser, localize light at specific wavelengths. Applied accelerations that alter the slot width also alter the photonic bandgap.

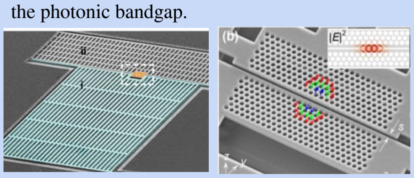


*Figure 2: (Left) Motional mass in blue. Photonic crystal (orange) is placed between the two masses. (Right) E-field irradiance shown within photonic crystal, with displaced holes marked with arrows.*

- Proof mass is driven by **radiation pressure** (i.e. force due to lasers).
- Mass motion alters resonant wavelengths of guiding laser. These shifts show up as dips in optical transmission and changes in physical oscillation frequency.

## Laboratory Setup

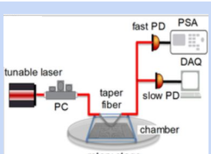
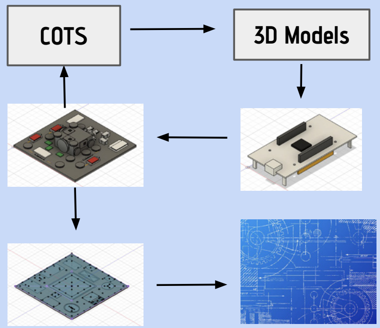



*Figure 3: Laser passes through a polarization controller, then the accelerometer, and optical transmission is monitored.*

- Rotating stage allows for known accelerations to be applied.
- Two photodetectors help detect optical transmission. Mechanical resonance frequency is measured simultaneously.

## Materials & Methods




- The first step involves the design of the IMU and its integration into a compact payload suitable for a space vehicle. Mechanical components that match required specs need to be found from off-the-shelf parts (COTS).
- Components need to be arranged in an assembly following connection and thermal considerations.
- Appropriate mechanical components need to be placed to mount the required parts. Correct identification of screws, spacers, spools and other components is key for mission success and to avoid mechanical failure.
- A finalized CAD model is converted into a blueprint drawing, which is sent to a machine shop for fabrication.

## Results

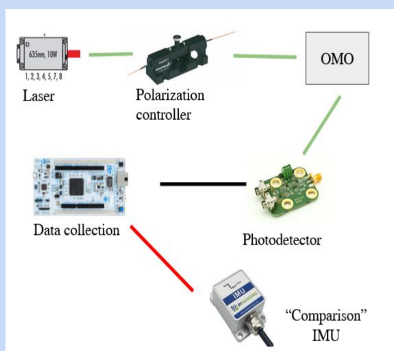



*Figure 4: The ORION laser routes through a polarization controller to the accelerometer, and optical transmission through the fibers is monitored by a photodetector. Data collection will compare the readout against that of conventional accelerometers, housed in a separate IMU module.*

- Three instances of the optical arrangement are needed, one per coordinate axis. These and the reference IMUs must be connected to the data collection via additional circuits.

## Results cont.

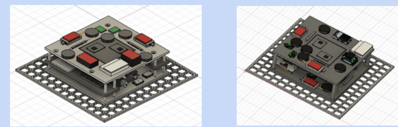


*Figure 7: (Left) Early draft of assembly, contains three levels to mount parts. (Right) Final assembly, minus half-plates out to the sides, contains extra components to complete necessary electrical connections*

- With each design, I aimed to reduce the effective area, as well as the weight of the system, without compromising any of the connections.

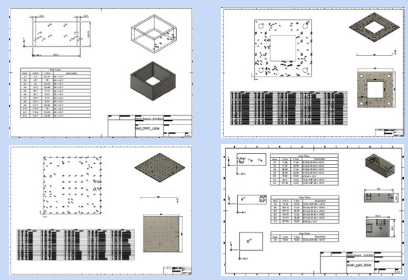


*Figure 8: Final blueprints*

- Blueprints made for each aluminum plate to be drilled, so that components can be attached and the complete assembly may be tested.

**Next steps:**

1. Machine the boards to contain all the holes for the connections
2. Assemble and test the configuration under a range of temperature, pressure conditions

## References

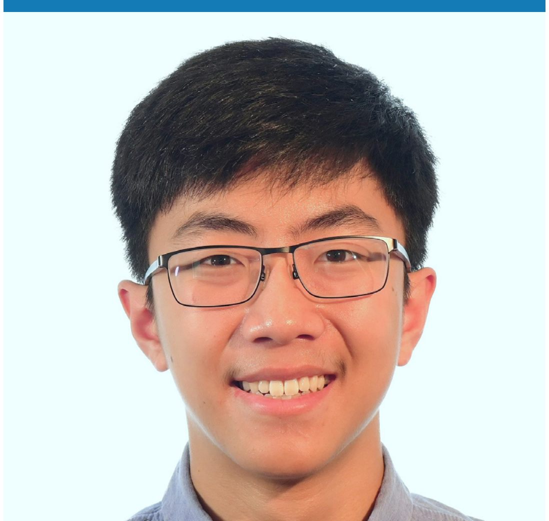
Flor, Jaime, Huang, Yongjun, et al. "A Chip‑scale Oscillation‑mode Optomechanical ... - Wiley Online Library." Wiley Online Library, 8 Apr. 2020, onlinelibrary.wiley.com/doi/abs/10.1002/lpor.201800329.

## Acknowledgements

I would like to thank Dr. Jaime Flor, Professor Chee Wei Wong, and the members of the Mesoscopic Optics and Quantum Electronics Laboratory at UCLA for their ongoing support in this project.


## ALBERT DONG

Computer Engineering
1st Year, UCLA


## CHARLES ZHANG

Electrical and Computer Engineering
1st Year, UCLA



**FACULTY ADVISOR**
**Ankur Mehta**

**DAILY LAB SUPERVISOR**
**Ankur Mehta**

**DEPARTMENT**
**Electrical and Computer Engineering**


# Computational Design of Bead-Based Wire Jamming Mechanisms


## ABSTRACT

When a wire is strung through a set of beads, applying tension to the wire compresses the beads together, increasing the rigidity of the structure. These bead-based wire jamming structures enable applications where a controllable level of stiffness is a key functionality, like deformable grippers. We aim to accelerate the design and research of these structures by developing software that allows the user to computationally generate 3D models of a large set of beads, eliminating the usage of CAD software.

We created a web-app that supports the ability to quickly preview, modify, and export 3D bead models based on user-inputted numerical parameters defining the desired geometry of each bead. The app accepts parameters of the size, length, and interface angle/shape of each bead, and supports the generation of cylindrical and spherical beads as well as multi-bead line segments and arcs. Using these software tools and a resin 3D printer, we generated, 3D printed, assembled, and iterated through multiple versions of various bead structures. By applying several bead geometries and performing static analysis, we designed self-deploying polygonal structures with multi-angled joints, as well as shape-shifting structures controlled by running multiple wires through each bead. Furthermore, we implemented the generation of both polygonal and shape-shifting structures in our web-app, increasing the design space it supports. The capabilities introduced in our research provide a higher-level method of designing beads, accelerating the design work necessary to investigate the properties of complex bead structures and their applications in soft robotics and self-deploying structures.

# Computational Design of Bead-Based Wire Jamming Structures


## Albert Dong, Charles Zhang, Prof. Ankur Mehta

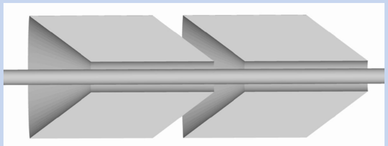
*Department of Electrical and Computer Engineering, University of California – Los Angeles*


UCLA Samueli School of Engineering

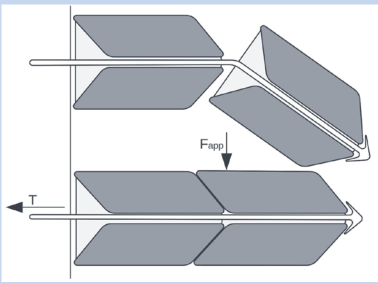
SUMMER UNDERGRADUATE RESEARCH PROGRAM

LEMUR

## Bead-Based Wire Jamming Structures



- Bead-based wire jamming structures are built by stringing a wire through a set of beads
- Cutaway shows beads with conical interfaces




- Applying tension (T) assembles the bead structure together
- Tightly interlocked beads lead to greater structural stiffness when resisting applied force ($F_{app}$)
- Designing structures using differing bead geometries leads to varying properties when tension is applied

## Objective

We aim to accelerate the design and research of bead-based wire jamming structures by developing software that allows the user to computationally generate 3D models of a large set of beads, eliminating the usage of CAD software.

## Methods

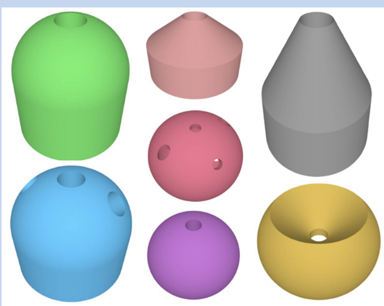
I. Create program that generates 3D models of beads given user specifications

II. Link bead generation program to user interface to enable preview and iteration

III. Generate/prototype beads and use them to create bead structures

IV. Implement generation of the investigated bead structures

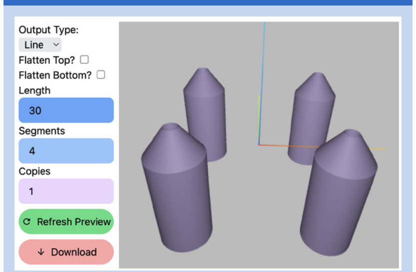
## Bead Geometries



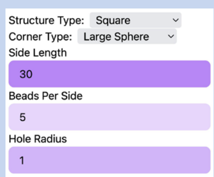
As examples, in left to right, top to bottom order:

- Cylindrical bead with spherical top
- Cylindrical beads with conical top (x2, different sizes and interface shapes)
- Spherical bead with 3 wire holes/paths
- Cylindrical bead with spherical top, 3 wire holes/paths
- Spherical bead (simple)
- Spherical bead with conical cutout to allow wire to exit bead at a range of angles

## User Interface

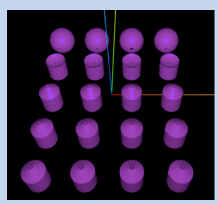



## Design and Manufacturing



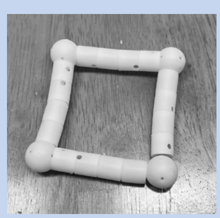

Geometric parameters of a bead/bead structure are inputted into web-app user interface

3D models of beads are generated and displayed for review



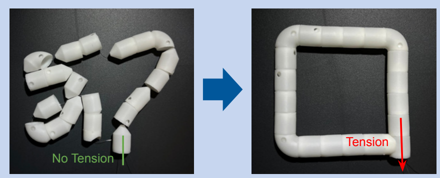
3D models are uploaded to resin 3D printer and manufactured

Beads are assembled into structure by stringing a wire through holes in beads

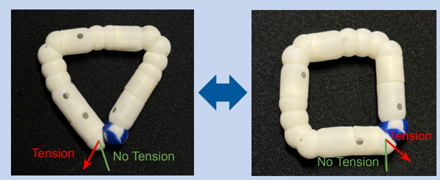


## Sample Bead Structures

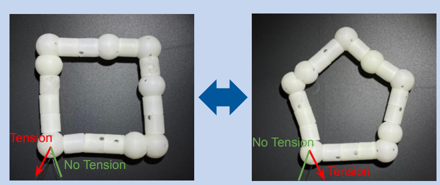
- Presented samples are generated using web app and printed on Formlabs Form 3




Single-wire square bead structure with corners formed by curved, cylindrical beads that deploys via wire tension




Two-wire bead structure with multi-bead spherical joints shifting between triangle and square shapes when tension is applied through one of the wires running through the structure



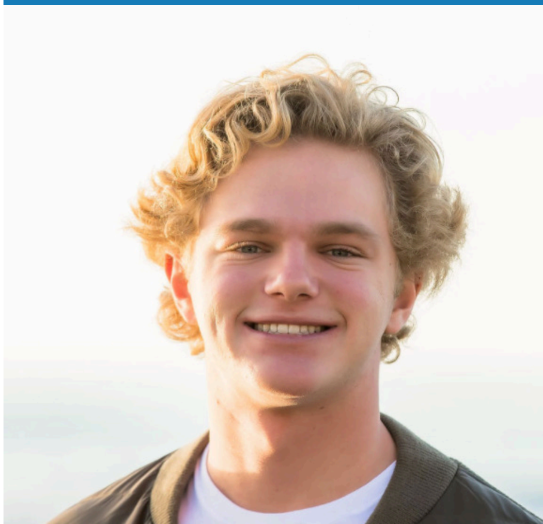

Two-wire bead structure with single-bead spherical joints shifting between square and pentagon shapes

## Conclusion and Future Prospects

- Our research enables high-level design of bead-based wire jamming structures
- Current web-app can be extended with additional bead designs and support for more complex generated structures
- Design tools can be used to research applications in soft robotics and deployable structures


## Acknowledgements

We would like to thank Professor Ankur Mehta and the members of the Laboratory for Embedded Machines and Ubiquitous Robots (LEMUR) for their continued guidance, as well as the SURP and Fast Track programs for offering us the opportunity to do research this summer!




**FACULTY ADVISOR**

Greg Pottie

**DAILY LAB SUPERVISOR**

Jeffery Jiang

**DEPARTMENT**

Electrical and Computer Engineering

## ALEX BRANCH

Electrical and Computer Engineering
2nd Year, UCLA


# Uncertainty Thresholded Loss Optimization

## ABSTRACT


We present a novel loss modification and training technique called Uncertainty Thresholded Loss Optimization (ThreshOptim) designed to enhance the generalization performance and training computation of classification models for image and tabular data. Our proposed method addresses the challenge of Out-of-Distribution (OOD) generalization, where models tend to struggle when presented with samples from unseen distributions. ThreshOptim is a thresholding indicator function during training, which selectively drops certain samples in a given batch per training epoch if they exceed a predefined softmax probability.

This thresholding approach ensures model capacity in training is dedicated to the highest entropy samples. By training on these selected sample, our model trains on the hardest points -- points we suspect lead to better regularization -- while ignoring the easiest points, improving OOD generalization and maintaining competitive performance on in-distribution data. Through extensive experiments on benchmark datasets and model architectures, we demonstrate the effectiveness of ThreshOptim in maintaining competitive in-distribution performance. Specifically, its ability to retain testing accuracy while training on a small subset of the dataset. We highlight the potential of our method to address the challenging problem of generalizing well to novel and unseen data distributions with an easy-to-implement change, making it a valuable addition to the arsenal of techniques for enhancing model robustness and adaptability in real-world applications.

# Uncertainty Thresholded Loss Optimization (ThreshOptim) To Enhance Model Generalization

Alexander Branch
Prof. Gregory Pottie
Department of Electrical and Computer Engineering, University of California – Los Angeles

UCLA Samueli School of Engineering

SUMMER UNDERGRADUATE RESEARCH PROGRAM

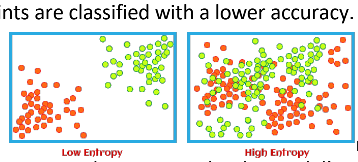
## Introduction

In machine learning, standard practice is to train your model on the full dataset each epoch. Some samples perform better than others, and the model adjusts its weights. Previous works have proven that certain points are "more important" in training than others. [1]

Our method aims to focus our model to train on these "important" points. We propose ThreshOptim which calculates a model's performance on a sample and selectively trains on the model based on a predefined threshold. With ThreshOptim, our model prioritizes training on the "hard" samples, forgoing the "easy" points which we classify with high accuracy.

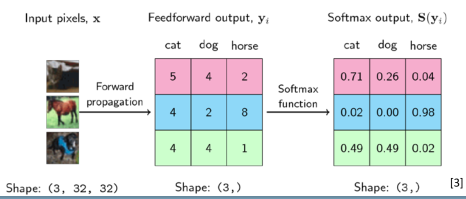
## Background

Selects points with high entropy -- measure of disorder. High entropy points are classified with a lower accuracy.




[2]

Softmax function used to convert the the model's predictions into a probability distribution (0 -> 1).



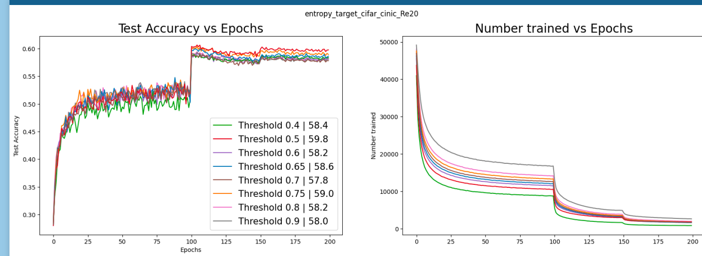

[3]

## Methods

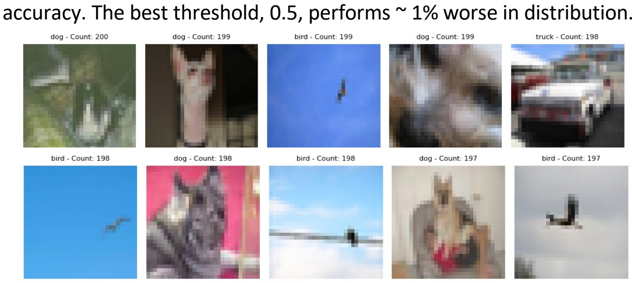
1. Initialize the model -- ResNet 20.
2. Load CIFAR-10 dataset for training. [4]
   - 50,000 images
   - 10 classes


3. Load CINIC-10 dataset for testing. [5]
   - 90,000 images (Out-of-distribution)
   - Same 10 classes
4. Forward pass dataset through model calculating entropies and SoftMax.
5. Remove samples from dataset that outperform threshold.
6. Backpropagate on this new dataset.
7. Repeat from step 4 with the the full dataset until 200 epochs.

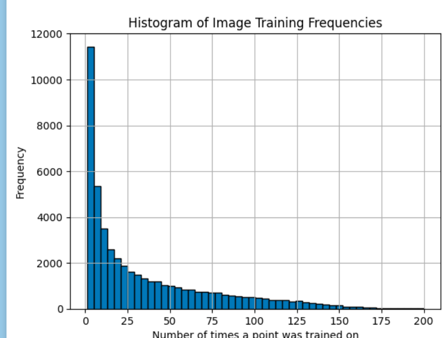
## Results and Discussion




ThreshOptim maintains the standard SGD baseline of ~ 60% OOD accuracy. The best threshold, 0.5, performs ~ 1% worse in distribution.




The images ThreshOptim selects most are challenging with the model's prediction often falling split between two classes.




ThreshOptim trains on 16% of the samples of a normal model. Most points are trained on less than 10 times, making only the "hard" samples (image above) see repeated training. A few samples are trained on each epoch

## Conclusion

ThreshOptim shows comparable OOD accuracies while remaining competitive in distribution. This is all done while training on a small subset of the original dataset – [end training percent] by the last epoch. With proper setup, this means we will be able to train much faster by using a smaller dataset, yet still maintain the accuracies you would expect without ThreshOptim.

## Acknowledgements

[1] An Empirical Study of Example Forgetting during Deep Neural Network Learning, Mariya Toneva. arXiv:1812.05159
[2] Understanding Entropy: the Golden Measurement of Machine Learning, Andrew Ye. Towards Data Science
[3] You Don't Really Know Softmax, Sewade Ogun.
[4] Learning Multiple Layers of Features from Tiny Images, Alex Krizhevsky, 2009. cs.toronto.edu
[5] Group Knowledge Transfer: Federated Learning of Large CNNs at the Edge, Chaoyang He arXiv:2007.14513
Thanks to the Interactive Systems REU for funding our research this Summer.
Thanks to Sunay Bhat for his idea of, and initial work on, ThreshOptim


## ALEXANDER HENDERSON

Physics
1st Year, UCLA


**FACULTY ADVISOR**

**Lihua Jin**

**DAILY LAB SUPERVISOR**

**Boliang Wu**

**DEPARTMENT**

**Mechanical and Aerospace Engineering**

# Improved Response of Shape Memory Polymer Composites Using Eutectic Gallium Indium


## ABSTRACT

Shape memory polymers (SMPs) are polymers characterized by their intrinsic ability to “remember” their original shape. Able to hold a temporary shape when deformed and cooled, these polymers will revert to their original configuration upon exceeding a glass transition temperature. When exposed to a temperature gradient, SMPs exhibit a bending phenomenon due to non-uniform strain within the material. These unique properties invite exciting applications for biomechanical devices and soft robotics. While the actuation mechanism of SMPs is primarily governed by heat transfer and viscoelastic relaxation, a key limitation is their intrinsically low thermal conductivity, resulting in traditional SMPs exhibiting slow response speeds.

To enhance the performance of these polymers, our research integrated SMPs with liquid metal to improve thermal conductivity and, by extension, response speed. By creating a roughly even suspension of an eutectic alloy of gallium and indium (EGaIn) within a polymer matrix using a mortar and pestle, we combined the attributes of SMPs and the exceptional conductivity of EGaIn. Preliminary results indicate that this composite material exhibits greatly improved thermal conductivity and sports actuation response times that are 100% quicker compared to SMPs devoid of liquid metal. Additionally, we present several interesting mechanisms that take advantage of this improved response time to enable the material to achieve temporary locomotion and snappy jumps. The goal of this work is to establish a methodology for fabricating SMPs with quick response times for applications in the fields of soft robotics and biomechanical devices.

# Improved Response of Shape Memory Polymer Composites using Eutectic Gallium Indium

UCLA Samueli School of Engineering
SUMMER UNDERGRADUATE RESEARCH PROGRAM

Alexander Henderson[1], Boliang Wu[1], Lihua Jin[1] (PI)

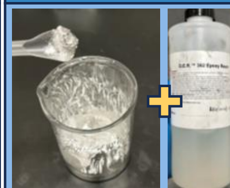
1. University of California, Los Angeles, Department of Mechanical and Aerospace Engineering

## Introduction

Stimuli-responsive materials are materials that react to certain environmental changes, such as temperature, electric and magnetic fields, and humidity, and promise many applications as deformable actuators. One such material is the shape memory polymer (SMP), which is capable of being deformed, holding that deformation, and then returning to its original configuration when heated. However, this recovery process is quite slow as it is limited by an intrinsically low thermal conductivity.

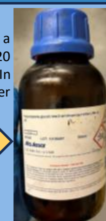
Other groups have managed to increase the conductivity of different polymer-based materials by embedding liquid metal within polymers and elastomers. Taking inspiration from these works, we have created a suspension of Eutectic Gallium-Indium (EGaIn) in an SMP to greatly increase its recovery speed while still retaining its flexible properties. After much trial and error, we have developed a methodology for preparing SMPs with a consistent distribution of EGaIn that display increased thermal conductivity and response times.

## Materials and Methods

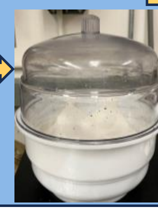


Step 1: MIX
EGaIn is prepared by combining gallium and indium (75% gallium : 25% indium by weight) on 200C hotplate for 24 hours and then stirring until homogenous.
EGaIn and Epoxy Resin are combined in a 70%:30% EGaIn:Epoxy volume ratio.

Step 2: MASH
Solution is mixed with a mortar and pestle for 20 min, or until sizable EGaIn particles are no longer visible in the solution.
The solution is then transferred to a glass vial.



Step 3: ADD
The curing agent then is added in a 1.25:1 Epoxy:Agent weight ratio.
The solution is then thoroughly mixed at for 5 min.



Steps 4 and 5: PREPARE
Sample is placed in vacuum chamber to remove air bubbles for 1 hour. Sample is moved to 45C oven for 2 hours and 20 min to thicken.
Sample is stirred to redistribute EGaIn and then transferred into petri dish for curing.

Step 6: CURE
Sample is cured in 60C oven for 8+ hours

## Results

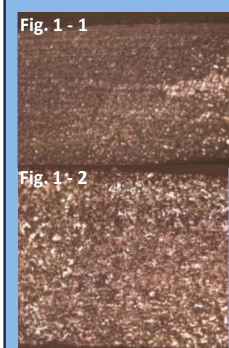
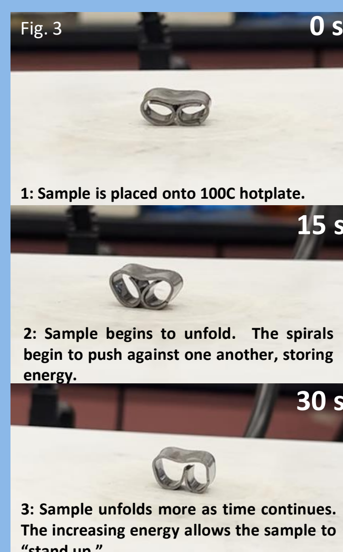



Fig. 1-1, 1-2: Cross sections of EGaIn SMP under 5X(top) and 10X(bottom) magnification. EGaIn particles average 10-50um. This even distribution leads to increased thermal conductivity.

Fig. 2: The recovery speed of the SMP with EGaIn (left) is twice as fast as the one without (right). Both samples are placed into a 100C incubation chamber.

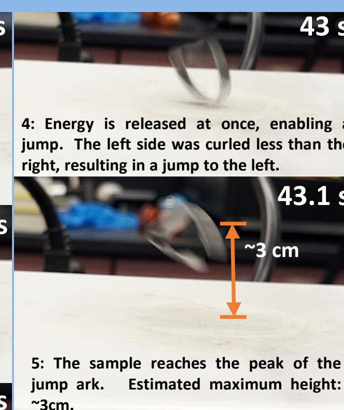




Fig. 3: A EGaIn SMP placed into a 100C hotplate. The dual spiral geometry allows the sample to contact itself and store energy. After a certain threshold, this stored energy is quickly released, resulting in a sizable jump. Additionally, by unevenly loading this sample, we can control the direction of the locomotion.

## Conclusion and Future Works

**Overall, our compound of EGaIn within an SMP demonstrates…**

- 100% faster response times than samples without liquid metal
- Higher thermal conductivity

**We have determined that the jumping mechanism is due to…**

- The SMP coming into contact with itself
- Friction allows it to store energy in the form of strain
- Releasing this energy all at once enables "snappy" jumps

**Interestingly…**

- SMP's without liquid metal cannot jump - they simply unfold
- The lower thermal conductivity prevents the strain from distributing throughout the material
- The more-concentrated strain snaps the material

**Some future works may include…**

- Working towards even faster response times
- Instant snap back to the original shape when heated

## Acknowledgements

I would like to thank Professor Lihua Jin, Boliang Wu, and the Mechanics of Soft Materials Laboratory for making this project possible.

Additionally, I would like to thank the SURP program, William Herrera, and the SURP facilitators for making this opportunity available to countless students.

This work is supported by the National Science Foundation (NSF) through Grant No. CMMI-2048219 and a NSF Research Experiences for Undergraduate (REU) supplement Grant No. CMMI-2328091.

## References

1. Hongzhang Wang, Youyou Yao, Xiangjiang Wang, Lei Sheng, Xiao-Hu Yang, Yuntao Cui, Pengju Zhang, Wei Rao, Rui Guo, Shuting Liang, Weiwei Wu, Jing Liu, and Zhi-Zhu He, Large-Magnitude Transformable Liquid-Metal Composites. *ACS Omega* 2019 *4* (1), 2311-2319

2. Jinwei Cao, Xin Li, Yiwei Liu, Guang Zhu,Run-Wei Li, Liquid Metal-Based Electronics for On-Skin Healthcare, Biosensors, 13, 1, (84), (2023). https://doi.org/10.3390/bios13010084

3. Wang, Q., Yu, Y. and Liu, J. (2018), Preparations, Characteristics and Applications of the Functional Liquid Metal Materials. Adv. Eng. Mater., 20: 1700781. https://doi.org/10.1002/adem.201700781


**ANGELIQUE JAYADINATA**

Materials Engineering
3rd Year, UCLA


**FACULTY ADVISOR**

**Laura Kim**

**DAILY LAB SUPERVISOR**

**Laura Kim**

**DEPARTMENT**

**Materials Science and Engineering**

## Optimization of Light-Trapping Diamond Waveguide for Low-Power-Consumption Diamond Sensing Platform


### ABSTRACT

Nitrogen vacancy (NV) centers in diamond have emerged as one of the leading solid-state quantum systems. Their optically addressable spin states make the NV system appealing for magnetic field sensing and imaging applications in condensed matter physics, neuroscience, geophysics, and device analysis. However, the low absorption cross section of NV centers limit the conversion of excitation optical power to spin-dependent fluorescent signals, leading to high optical power consumption, therefore precluding applications where phototoxicity and heating effects are of concern. This issue can be addressed with diamond light-trapping waveguide structures, which have shown to increase optical path length via total internal reflection. Using a ray tracing simulation tool, I optimized tunable parameters such as beam incidence angle, beam incidence position, and facet geometry of a diamond waveguide to maximize optical path length.

The results show an optical path length of up to 90 cm in a 3mm x 3mm sample. This represents an enhancement factor of nearly 300 in pump-to-signal photon conversion efficiency compared to a single pass geometry. The coupling of a green excitation beam into the waveguide structure is demonstrated with a precision experimental apparatus that allows translational and rotational degrees of freedom. I show that the mechanical polishing of a diamond waveguide minimizes scattering losses and enables us to achieve the computationally predicted optical path lengths. Its potential to address a greater number of NVs within a given excitation power budget demonstrates a promising future for a compact and portable precision diamond sensing platform.

# Optimization of Light Trapping Diamond Waveguide for Low-Power Magnetic Field Sensing

Angelique Jayadinata, Prof. Laura Kim
*Department of Materials Science and Engineering, University of California - Los Angeles*

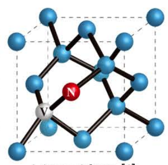
UCLA Samueli School of Engineering
SUMMER UNDERGRADUATE RESEARCH PROGRAM

Samueli Research Scholars UCLA Samueli School of Engineering

## Motivation & Theory

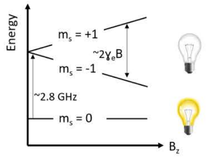
Nitrogen Vacancy (NV) Centers in Diamond

NV centers are defects in diamond lattice consisting of a substitutional nitrogen atom adjacent to a vacant site. Both its optically sensitive spin state energy levels and its long coherence time at ambient temperatures make NV centers appealing for magnetic field sensing and imaging applications.

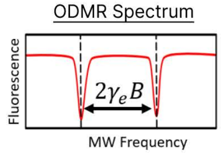



Adopted from [1].

Working Principle of NV Magnetometers



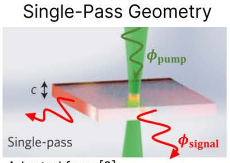

ODMR Spectrum




NV centers have a spin triplet electronic ground state with a zero field splitting of approximately 2.8 GHz. The degeneracy of $m_s = \pm 1$ sublevels is lifted by an external magnetic field via Zeeman splitting. The energy splitting can be measured by optically detected magnetic resonance (ODMR).

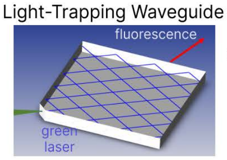
Pump-to-Signal Photon Conversion Efficiency

Single-Pass Geometry




Adopted from [2].

Light-Trapping Waveguide




Pump-to-signal photon conversion efficiency:

$$\eta_{ODMR} = \frac{\phi_{signal}}{\phi_{pump}}$$

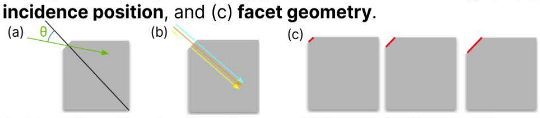
The low absorption cross section of NV centers results in a low pump-to-signal photon conversion efficiency in a single pass geometry, leading to high optical power consumption.

Clevenson et al.[2] demonstrated a diamond waveguide structure increases optical path length via total internal reflection. Such strategy is implemented to improve the pump-to-signal conversion efficiency.

[1] Ishii S. et al. Quantum Beam Sci. 6(1), (2) (2022); [2] Clevenson, H. et al. Nature Physics 11(5) (2014)

## Objectives

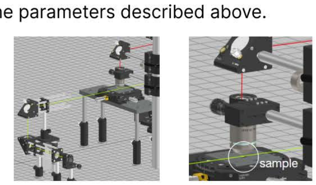
- Maximize the optical path length in a light trapping diamond waveguide by optimizing (a) **beam incidence angle**, (b) **beam incidence position**, and (c) **facet geometry**.




- Achieve tunable density of rays within a selective sensing region of interest.

## Experimental Methods

- Run non-sequential light propagation simulations to calculate optical path lengths and optimize the parameters described above.
- Construct an optical experimental set up with translational and rotational degrees of freedom to implement optimized excitation conditions.

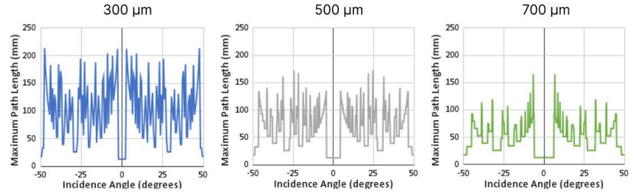


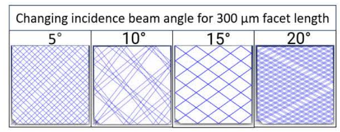

- Collect fluorescence signals and compare corresponding pump-to-signal photon conversion efficiencies of the light trapping diamond waveguide to that of a single pass geometry.

## Results

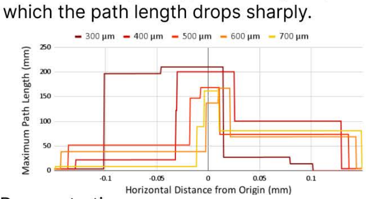
Simulation Results

- Shorter facet lengths result in greater path lengths with significant fluctuations, while longer facet lengths yield more stable, but shorter path lengths due to the increased probability for light to couple out of the sample.

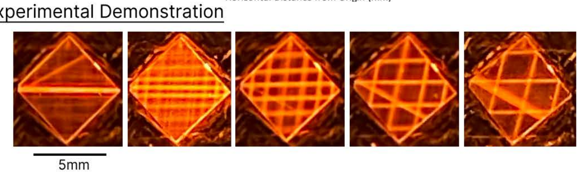



- Sweeping through the horizontal beam incidence position determined that the maximum path lengths were at the origin, resulting in stable values up to approximately 45 µm away from the origin, after which the path length drops sharply.




Experimental Demonstration




- Red fluorescence images showing the total internal reflection pathway inside a 5mmx5mm diamond sample captured using a red filter. The incidence light comes in from the left, where the 500µm input facet is normal to the beam in the left-most picture.

## Conclusion

- The simulation results show an optical path length of up to 90 cm in a 3mm x 3mm diamond substrate. This represents an enhancement factor of nearly 300 in its pump-to-signal photon conversion efficiency compared to the single pass geometry.
- Its potential to address a greater number of NVs within a given excitation power budget demonstrates a promising future for a compact and portable precision diamond sensing platform.

## Future Work

Incorporate resonant nanophotonic structures to improve the interaction time of NV centers with the incoming excitation light, thus leading to a further reduction of power consumption of diamond quantum sensors.

## Acknowledgements

I would like to extend my sincere gratitude for my PI, Dr. Laura Kim, for her continuous support and mentorship throughout the program, and my colleague, Alex Jensen, for his help this summer. Additionally, I would like to thank the entire UCLA SURP facilitators for making this program possible, as well as the SRS for funding this research.


FACULTY ADVISOR

**Richard Wesel**

DAILY LAB SUPERVISOR

**Tim Jacques**

DEPARTMENT

**Electrical and Computer Engineering**

**ARSHNOOR DAHLIWAL**

Electrical and Computer Engineering
4th Year, UCLA


## Analysis of Min-Star Decoder for LDPC Codewords of Different Code Rates


### ABSTRACT

Information transmitted over a noisy channel can be corrupted by noise that can flip or erase bits. One technique to recover the original signal is to send an encoded message that carries redundant bits of information, known as parity bits. Low Density Parity Check (LDPC) codes are widely used for encoding messages and introducing parity bits in many forms of digital communication. Their key feature is a parity-check matrix with a low density of ones, which allows for efficient decoding by requiring a small number of computations per parity check equation. There are several algorithms available to decode LDPC codes via iterative message passing to estimate the message bits. Min-sum is one such algorithm, which approximates the theoretical node computations by simple minimum and addition operations. The Min-star algorithm reduces frame error rate (FER) with minimal additional complexity by adding a correction factor to the determination of the minimum. In this research, we investigate the accuracy of Min-Star compared to Min-Sum and normalized Min-Sum for different code rates, a ratio that determines the amount of redundancy added to a message.

To compare their correction capacities, FER curves are plotted against signal-to-noise ratio (SNR). Our results show that Min-Star results in lower FER than Min-Sum and Norm-min when all algorithms are limited to 100 iterations. Furthermore, we adapt an FPGA implementation of Min-Star to work with the different code rates that match the performance standards set by the Consultative Committee For Space Data Systems (CCSDS). Similarly, we generate FER plots for different Code Rates and plot them against the CCSDS performance curves. We find that our implementation performs in alignment with the standard.

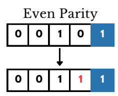
# Optimizing an FPGA implementation of Min-Star decoding of LDPC codes

UCLA Samueli School of Engineering
SUMMER UNDERGRADUATE RESEARCH PROGRAM

Arshnoor Dhaliwal, Timothy Jacques, Egor Glukhov, Matheo Irazabal, Dr. Richard Wesel

## BACKGROUND

- Channel **noise** inevitably leads to bit flips and erasures
- Parity Bits are added to store redundant information about a message before transmission
- When parity constraints fail, they may indicate the errors present




## LDPC

- Low-Density Parity Check (LDPC) codes are used to correct errors in wireless, optical, and satellite communication standards
- The low density of ones in the parity matrix reduces the decoder complexity

## Parity-check (H-)matrix

- Defines constraints that must be satisfied by a valid codeword.
- Each row is a parity-check equation that represents connections between parity bits and the original message
- Below is an example of a high-density H-matrix

H-matrix with Codeword
H ⊕ C = 0

$$H = \begin{bmatrix} 0 & 1 & 0 & 1 & 1 & 0 & 0 & 1 \\ 1 & 1 & 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 & 1 & 0 & 1 & 0 \end{bmatrix}$$

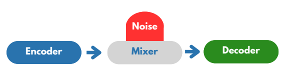
Top row Parity-Check Eqn:
C2 ⊕ C4 ⊕ C5 ⊕ C8 = 0

## Generator Matrix

- A generator matrix is used to encode the message and introduce parity bits
- The message (a row vector) is encoded by right-multiplying with the generator matrix to produce the codeword.
- The codeword is then transmitted over channel.

## Decoder

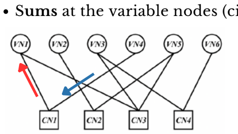
- An H-matrix can only validate if a codeword is valid, not correct for errors.
- Therefore, a **decoding algorithm** must be used to transform the noisy codeword that was received into a valid, noiseless codeword.




- A popular algorithm is Belief Propagation (BP)
  - decoding involves iterative message passing (nodes talk back and forth until a valid codeword has emerged)
  - balance between robustness and computational complexity

## Min-Sum / Min-Star

- **Min-Sum**: an iterative decoding algorithm that performs two computations:
- **Minimums** at the check nodes (squares)
- **Sums** at the variable nodes (circles)




Connections between nodes described by H-matrix

$$u^{(t)}_{c_i \to v_j} = \beta^{(t)}_{(c_i, v_j)} \times \prod_{v_{j'} \in \mathcal{N}(c_i)/\{v_j\}} \mathrm{sgn}(l^{(t-1)}_{v_{j'} \to c_i}) \times \min_{v_{j'} \in \mathcal{N}(c_i)/\{v_j\}} \left| (l^{(t-1)}_{v_{j'} \to c_i}) \right|$$

$$l^{(t)}_{v_j \to c_i} = l^{ch}_{v_i} + \sum_{c_{i'} \in \mathcal{N}(v_j)/\{c_i\}} u^{(t)}_{c_{i'} \to v_j}$$

- Instead of passing bits, the algorithm passes **LLRs** (Log Likelihood Ratios) which represent the probability that a bit is 1 or 0.

---

- After many iterations, the LLRs diverge to ±∞, indicating high confidence in a certain bit value.
- An improved algorithm, **Min-Star**, adds a correction factor to the sum portion to more accurately approximate to an LLR calculation.

$$l^{(t)}_{v_j \to c_i} = l^{ch}_{v_i} + \sum_{c_{i'} \in \mathcal{N}(v_j)/\{c_i\}} u^{(t)}_{c_{i'} \to v_j}$$ **+** correction Factor

- Higher accuracy while retaining low complexity

## INTRODUCTION

The goals of this research are as follows:

- Compare the performance of the Min-Sum and Min-Star algorithms.
- Implement a Min-Star decoder for the CCSDS (Consultative Committe for Space Data Systems) encoding that works for code rates of 1/2, 2/3 and 4/5, allowing for customizability between throughput and robustness
- Explore various methods for performance optimization

## METHODS

**1** Follow the CCSDS guidelines to generate both the H-matrixes and Generator Matrixes for simulation

$$H_{1/2} = \begin{bmatrix} 0_M & 0_M & I_M & 0_M & I_M \oplus \Pi_1 \\ I_M & I_M & 0_M & I_M & \Pi_2 \oplus \Pi_3 \oplus \Pi_4 \\ I_M & \Pi_5 \oplus \Pi_6 & 0_M & \Pi_7 \oplus \Pi_8 & I_M \end{bmatrix}$$

$$\pi_k(i) = \frac{M}{4}\left(\left(\theta_k + \lfloor 4i/M \rfloor\right) \bmod 4\right) + \left(\phi_k\left(\lfloor 4i/M \rfloor, M\right) + i\right) \bmod \frac{M}{4}$$

CCSDS The Consultative Committee for Space Data Systems

- Each element is a circulant submatrix of size M
- Matrixes of H(2/3) and H(4/5) reuse the above matrix with additional columns, and smaller submatrix sizes

**2** Encoder-to-decoder Pipeline in MATLAB

- Write functions to encode a message-bit stream, send it through an additive white Gaussian-noise channel, and decode it using Min-Star decoding

**3** CUDA Frame-Error Rate Simulation

- NVIDIA CUDA (Compute Unified Device Architecture) to allow for GPU computation
- Frame-error rate (FER) for a particular SNR is found by running simulation on codewords until a threshold number of errors in the decoded messages are found
- Remove parity-check equations with singly connected variable nodes, and plot the average iterations until parity-check pass

Signal-to-noise: ↑SNR = ↓ NOISE

**4** Implement Multiple Coderates on FPGA

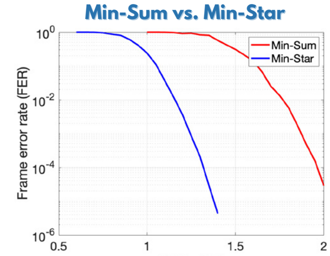
- Adapt an FPGA implementation of the Min-Star decoder for 2/3 and 4/5 rates (*in progress*)
- Write a top-level module for an encoder-mixer-decoder pipeline
- Tools: Xilinx Vivado and Vitis

## CONCLUSION

- Min-Star has significantly lower FER than Min-Sum
- Our implementation of the Min-Star for the different code rates closely matches the CCSDS performance standards
- Ignoring singly connected parity-check equations leads to lower FER for lower precisions

## RESULTS

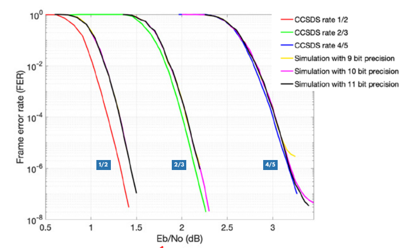
### Min-Sum vs. Min-Star




Eb = 1 ∴ Eb/No = SNR

Min-Star decoder has a lower FER than Min-Sum, indicating more accurate decoding. The difference in their FERs remains constant as the SNR changes.

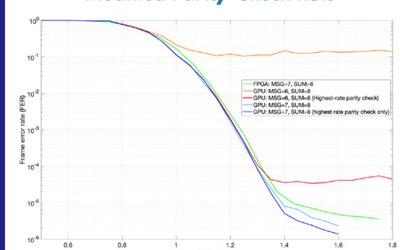
### Min-Star vs CCSDS Code Rates




*noise increases*

- Our implementations for all rates have FERs within 0.2 dB of the CCSDS's, with rate 1/2 having the highest FER difference
- Rate 4/5 has the same FER for SNR of 3 as rate 1/2 for SNR of 1.2

### Modified Parity-Check Rule




Ignoring singly connected variable nodes (highest rate parity check) leads to significantly better FER performance for lower precision of 6 msg bits and 8 sum bits. However, for higher precisions the performance difference becomes less significant.

## Acknowledgments

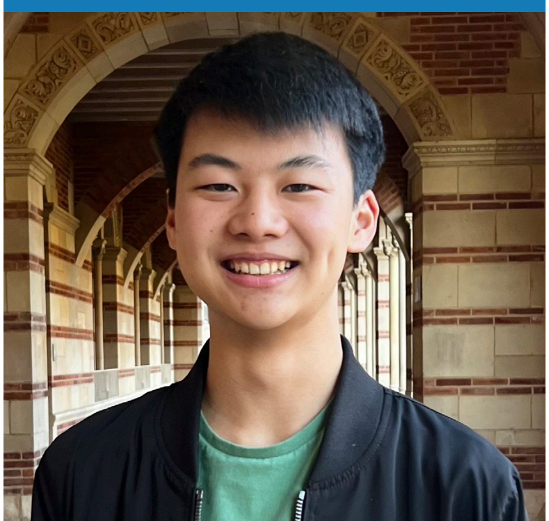
I would like to thank my team for their incredible support and work through this project: Timothy Jacques, Matheo Irazabal and Egor Glukhov. I would also like to thank Professor Wesel for this opportunity and his guidance.

This research is supported by the UCLA Communication Systems Laboratory, Lockheed Martin, and National Science Foundation Grant CCF-1955660.

*CCSDS Historical Document, public.ccsds.org/Pubs/131x1o2e2s.pdf. Accessed 8 Aug. 2023.*


## ARUNAN ELAMARAN

Electrical and Computer Engineering
1st Year, UCLA


## ETHAN LAI

Electrical and Computer Engineering
1st Year, UCLA



**FACULTY ADVISOR**

**Mani Srivastava**

**DAILY LAB SUPERVISOR**

**Pragya Sharma and Gaofeng Dong**

**DEPARTMENT**

**Electrical and Computer Engineering**


# Evaluating the Efficacy of Large Language Models as Frameworks for Personal Robotics

## ABSTRACT


Large language models (LLMs), or machine learning models specializing in processing and responding with human text, have shown the ability to reason with and parse complex tasks. Understanding such potential of LLMs, recent research has sought to integrate these models with the field of robotics to provide a means for users to interact more naturally with robots. In this project, we explore one implementation of an LLM-enabled system and the extent to which smaller, local LLMs can be used in edge-computing environments characteristic of personal robotic systems. Our system is distributed over a robot and a nearby edge server. The robot converts human-spoken instructions into a text prompt and sends it to a light-weight LLM on the edge server.

In response, the LLM generates code that calls custom function libraries; this script is sent back to the robot, which then utilizes its sensors and actuators to execute the user-specified task. We evaluate this system by prompting the LLM with tasks of varying complexity and ambiguity, measuring the latency and accuracy of each subprocess of the system's performance. Our findings indicate that the LLM consistently generates accurate code for simple, atomic tasks but occasionally generates erroneous code for complex tasks with higher levels of ambiguity. Despite these errors, we conclude that smaller versions of LLMs can be deployed effectively on edge-computing machines, but additional guardrails must be implemented to ensure their trustworthiness. Our framework serves as an initial platform for further testing of LLMs in robotics and other cyber-physical systems like home automation.

# Evaluating the Efficacy of Large Language Models as Frameworks for Personal Robotics

Arunan Elamaran, Ethan Lai
Collaborators: Yang Xing
Mentors: Pragya Sharma, Gaofeng Dong, Prof. Mani Srivastava

UCLA Samueli School of Engineering
SUMMER UNDERGRADUATE RESEARCH PROGRAM

## Background and Motivation

### Large Language Models

- Large language models (LLMs) are machine learning models specialized in understanding and generating human text.
- LLMs have shown promise in being capable of reasoning and task-parsing when given complex, abstract and physical tasks
- Resultantly, researchers are seeking to implement LLMs in the field of robotics so that humans may interact more naturally with these automated systems

### Goals

- Create a framework that incorporates an LLM into a robotic system to serve as the central reasoning node.
- Evaluate how LLMs work in real-time, resource-constrained environments that are often characteristic of home automation & personalized robotics.

## Methodology

### System Design

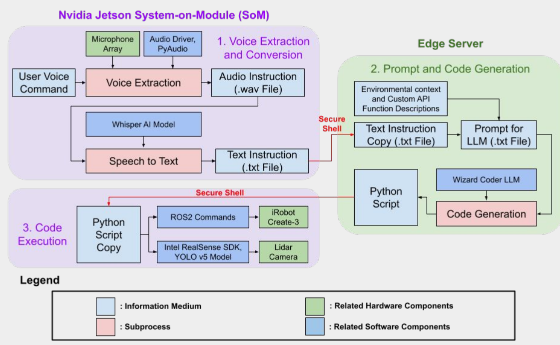



Figure 1: System Block Diagram

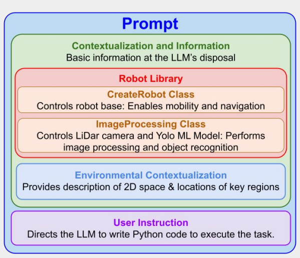



Figure 2: Prompt Structure Diagram

**WizardCoder LLM** was the code-generating LLM of choice. Crafting a descriptive yet succinct prompt was essential to the model yielding the desired output; overly comprehensive prompts negatively affected performance due to increasing token size. Determining the order of function composition (low vs. high level functions) was key to enabling LLMs to be flexible enough to produce code for any prompt while maintaining high accuracy and low latency.

### Testing

- 45 user provided tasks grouped into three categories (atomic, sequential, and composite) are passed into the framework
  - Atomic - Can be executed with single function call
  - Sequential - Must call multiple functions in succession, but without depth
  - Composite - Requires nested function calls based on environmental conditions
- Within each category, complexity for all the tasks is determined by the ambiguity of the instruction, thus allowing for evaluation of the LLM's reasoning abilities
- Latency and accuracy of each subprocess in the system is measured.
  - Latency is measured from the time the prompt is passed to the LLM until the output is stored in a variable.
  - Accuracy is measured by directional and syntactical correctness
    - Directional correctness - correct, necessary functions were called
    - Syntactical correctness - functions were called correctly

Github Project Page

## Results

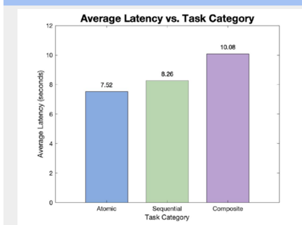



Figure 3: Plot of average latency for each task category. Latency increased by 9.8% for sequential tasks from atomic tasks and by 34% for composite tasks from atomic tasks.

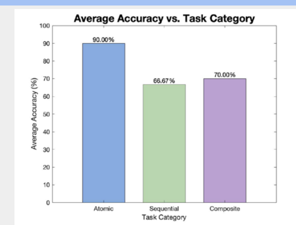



Figure 4: Plot of average accuracy for each task category. The system was less consistent when prompted with sequential and composite tasks, likely due to the fact that such tasks were generally more ambiguous and required stronger reasoning from the LLM.

An additional LLM was considered to parse the tasks and divide them into explicit subtasks. Optimal accuracy can be achieved by utilizing another LLM to create subtasks that can be individually passed into WizardCoder to produce code. However, this would only further increase the latency of the system.

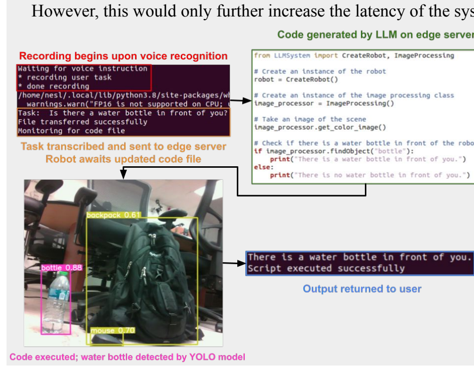



Figure 5: Example System Run. The LLM calls the correct functions to take a photo and run the object recognition model to find a water bottle.

When passed in other prompts such as "Find Jeff's water bottle", the LLM correctly reasoned that Jeff's items could be found at his desk and accordingly instructed the robot to go to his desk to search for his items. The coordinates of his desk (and other landmarks) are included in the environmental contextualization.

## Conclusion and Future Direction

LLM-enabled robotic systems showcase promising capability to accurately react to user instructions by producing correct code and determining the necessary steps to complete a task. However, since results indicate that the system is less consistent when prompted with more complex and ambiguous tasks, additional guardrails, such as prompt fine-tuning or more robust APIs, must be implemented to ensure its trustworthiness.

This system serves as an initial platform for further testing of LLMs and robotics. New modules and APIs may be added to customize robots or other localized systems for more specialized purposes. Further applications include cyber-physical systems like home automation.

## References

Vemprala, Sai, et al. *Chatgpt for Robotics: Design Principles and Model Abilities*, www.microsoft.com/en-us/research/uploads/prod/2023/02/ChatGPT___Robotics.pdf. Accessed 11 July 2023.

Shen, Yongliang, et al. "Hugginggpt: Solving AI Tasks with Chatgpt and Its Friends in Hugging Face." arXiv.Org, 25 May 2023, arxiv.org/abs/2303.17580.

## Acknowledgements

We would like to thank the UCLA Electrical and Computer Engineering Fast Track to Success Program and the Summer Undergraduate Research Program for this unique opportunity, as well as the Army Research Lab, Air Force Office of Scientific Research, and National Science Foundation for providing the NESL lab with funding to obtain the necessary project materials.


## ARYAK REKHI

Bioengineering
4th Year, UCLA

**FACULTY ADVISOR**
Dino Di Carlo

**DAILY LAB SUPERVISOR**
Cayden Williamson

**DEPARTMENT**
Bioengineering


# Analysis of Microbial Co-Cultures via PicoShells


## ABSTRACT

Current methods to characterize microbiomes, such as single-cell sequencing, do not allow for cell viability after the completion of an assay. In addition, separating and retaining the different microbes of a microbiome sample for analysis of growth and metabolite production remains a challenge; traditional methods such as well plate analysis are low-throughput and expensive. This issue is further complicated since these methods do not allow chemical signaling between colonies, which can further alter the phenotypes and expression profiles of the microbes in a sample. In this study, we investigate the growth of mixed microbial populations using the PicoShell platform. PicoShells are hollow, porous microparticles that encapsulate single cells and can be used to grow them into spatially discrete, clonal populations. Picoshells enable bulk suspension culture techniques to be used, maintaining an environment similar to the microbiome's native growth conditions.

To demonstrate the use of PicoShells for co-culture, we mixed Escherichia coli and Saccharomyces cerevisiae into one phase and encapsulated single cells into PicoShells. These shells were then cultured in different types of media, elucidating different relative growth behaviors. Through the use of flow cytometry, colonies were identified and sorted based on their cell type. This experiment demonstrates that multiple species can be grown together and separated for downstream analysis, acting as a proof of concept for the analysis of biological samples with large amounts of microbial diversity.

UCLA Samueli School of Engineering
SUMMER UNDERGRADUATE RESEARCH PROGRAM

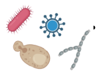
# Analysis of Microbial Co-Cultures via PicoShells

Aryak Rekhi ,Cayden Williamson, Dino Di Carlo

Department of Bioengineering, University of California Los Angeles, Los Angeles, CA, United States

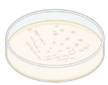
## Background

### Microbiomes



Analyzing growth and metabolite production of microbiomes is important for characterizing their interactions with the surrounding environment, such as the stomach or skin

### Current Limitations

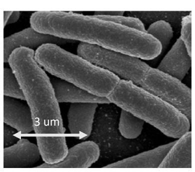


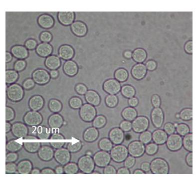
- Laboratory culturing conditions (agar plates, shake flasks) are not representative of the growth environments of microbiomes.
- Targeted colony picking from agar plates is low throughput (<2500 colonies/hour) and is limited to growth-based selection criteria.

### Questions

•Can two different cell types be screened and separated from a single population?
•Can interactions between colonies be studied through the PicoShell technology?

Bacteria (*Escherichia coli* ) Fungi (*Saccharomyces cerevisiae* )

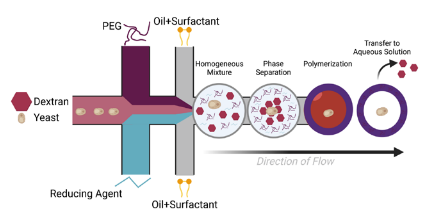


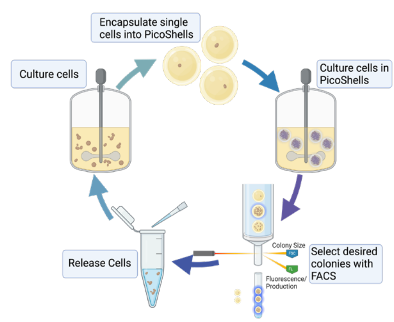

## PicoShells

- Polymeric, porous, hollow microparticles.
- Compartmentalize clonal subpopulations of yeast while enabling for waste and nutrient exchange.
- Compatible with FACS, allowing selection based on biomass (scatter), as well as metabolite production (fluorescence).

### Microparticle Fabrication




### PicoShell Workflow




## Conclusions

- PicoShells isolate monoclonal colonies while culturing microbes in bulk
- PicoShells made with multiple cell types remain stable
- Colonies can be sorted based on cell type through the use of side scatter
- Sorting efficiency exists in multiple media conditions, allowing for further analysis of growth patterns

- Next Steps: Acquire microbiome samples from clinic and test platform in comparison to other methods

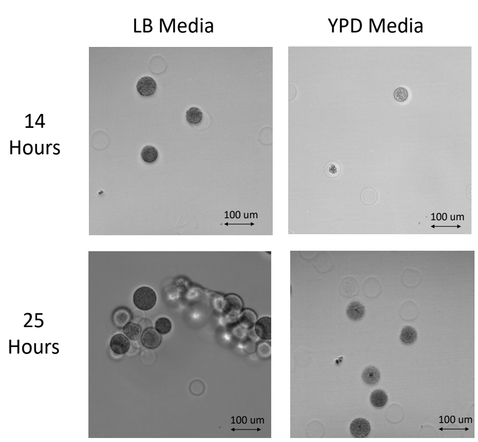
## References

- Van Zee et al., *PNAS* (2022)
- Cartoons made using BioRender
- "Escherichia Coli (E. Coli)." *Education*, education.nationalgeographic.org/resource/escherichia-coli-e-coli/. Accessed 22 Aug. 2023.
- "The Virtual Edge." *University of Wyoming*, www.uwyo.edu/virtual_edge/lab13/fungi_results.htm. Accessed 22 Aug. 2023.
- I would like to acknowledge the entire PicoPal team in the Di Carlo Lab for their support throughout the summer

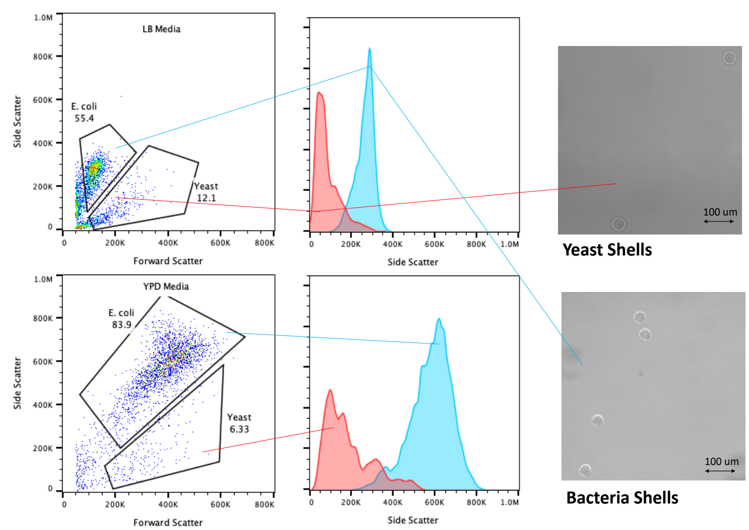
## Results

- Mixed *Escherichia coli* and *Saccharomyces cerevisiae* into dextran phase and fabricated PicoShells
- Grew PicoShell colonies in two different conditions (LB, 37$^0$C and YPD Media, 27$^0$C) for 25 hours
- FACS was used to identify and sort shells containing colonies with different microbial populations

### Images of Co-Culture PicoShells




FACS Scatterplots with Gated Populations | Side Scatter Histograms | Post-Sort Images of PicoShells




## BILL LI

Electrical and Computer Engineering
1st Year, UCLA


**FACULTY ADVISOR**
**Paulo Tabuada**

**DAILY LAB SUPERVISOR**
**Jonathan Bunton**

**DEPARTMENT**
**Electrical and Computer Engineering**

# LiDAR Point Cloud Registration Guarantees with PASTA Supervision


## ABSTRACT

LiDAR sensors are heavily adopted by both researchers and industry professionals for addressing localization challenges in autonomous systems. Localization for LiDAR requires aligning two LiDAR measurements obtained from different perspectives, occurring when the robot is in different positions and determining its movement. The alignment of the two scans or point clouds is referred to as the scan-matching or point-cloud registration problem. Existing algorithms for this problem are predominantly heuristic and local in nature, leading to inaccuracies when the initial alignment is suboptimal. Provably Accurate Simple Transformation Alignment (PASTA), designed by the UCLA CyPhy Lab, aims to provide formal error guarantees on localization error without relying on point-to-point correspondences as existing algorithms do. As most scan-matching algorithms utilize 3D point cloud data, we aim to determine the best-performing algorithms by comparing rotation and translation estimations with artificial 3D data generated from existing 2D LiDAR scans.

The nominal performance of several variations of such algorithms, including Iterative Closest Point (ICP), Fast Point Feature Histograms (FPFH), and PASTA, on 3D data was successfully validated through their estimated solutions to the manually applied rotation on the point clouds. Validation with ground-truth measurements for the alignment of LiDAR point clouds is in progress to determine the best algorithms for certain environments, in addition to attaching PASTA's guarantees to supervise alternative heuristic methods. Future work includes supervising machine learning point-registration with PASTA and integration into simultaneous localization and mapping (SLAM) frameworks for mapping of unknown environments and obstacle avoidance.

UCLA Samueli School of Engineering
SUMMER UNDERGRADUATE RESEARCH PROGRAM

# LiDAR Point Cloud Registration Guarantees with PASTA Supervision

**Bill Li, Jonathan Bunton, Matteo Marchi, and Professor Paulo Tabuada (Cyber-Physical Systems Lab)**
*Department of Electrical and Computer Engineering, University of California — Los Angeles*

## Background and Motivation

LiDAR sensors used for localization in autonomous systems – determining how a robot moves from measurements taken from different positions – require the alignment of two measurements obtained from different perspectives. This alignment of the two scans or point clouds is referred to as the **scan-matching** or **point-cloud registration problem**.

Existing scan-matching algorithms are predominantly heuristic and local in nature, and therefore do not return theoretical performance guarantees. CyPhy's Provably Accurate Simple Transformation Alignment (**PASTA**) solves this issue, providing **deterministic error guarantees** on localization error.

Previous testing was conducted with algorithms not designed for 2D data. More verification was thus required to determine their performance with 3D point clouds, which is more representative of more commonly used data.

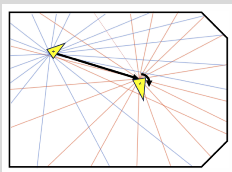


Figure 1: LiDAR scans taken from different positions (blue and red), showing how rays hit with different densities based on locations. Correspondences may not exist or be approximate, especially in further, lower-density areas.

## Results

The following graphs show pose estimation errors in the relative poses, quantified with SE(3) double geodesic distance, where two poses are represented as TA = (RA, tA) and TB = (RB, tB).

$$dist_g(T_A, T_b) = ||log(T_A^{-1}T_B)^\vee|| = ||log(T_B^{-1}T_A)^\vee||$$

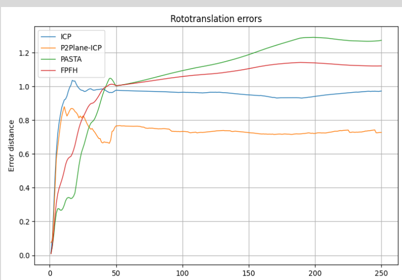



Figure 7: Roto-translational errors estimated by algorithms with pair-comparison lag. Higher lag indicates further physical distance (and a greater difference between point clouds) between poses.

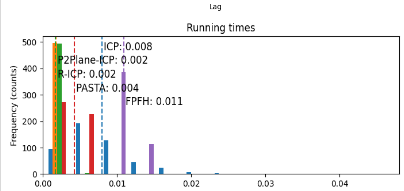



Figure 8: Average computed runtimes of various algorithms. Note PASTA's fairly low runtime performance.

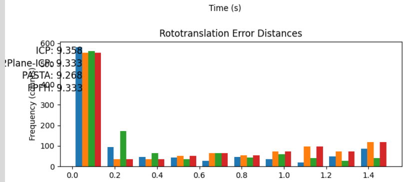



Figure 9: Roto-translational error distances vs frequency counts.

## Objective

We aim to determine the best-performing algorithms by comparing rotation and translation (roto-translation) estimations with simulated 3D data generated from existing 2D LiDAR scans.

## Materials and Methodology

**Materials**

- Robot with LiDAR scanner
- Python & additional packages (Scipy, Open3D, and others)

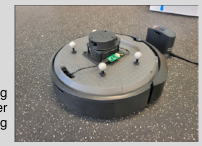


Figure 2: Robot moving with LiDAR scanner measuring surrounding environment

**PASTA Algorithm**

- Directly utilizing LiDAR point cloud data is unreliable. To increase robustness, registration is performed with the computed **convex hull.**

Figure 3: Example of point cloud, with surrounding convex hull as the smallest convex set containing all points

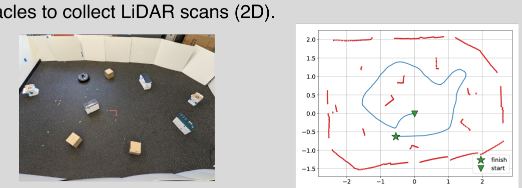
- By comparing the first **c** (centroid) and second **Σ** (covariance) moments of the hulls, the rigid transformation is found.

$$\mathbf{c} = \frac{\int_H \mathbf{x}\, \mathrm{d}\mu}{\int_H \mathrm{d}\mu}, \qquad \mathbf{\Sigma} = \frac{\int_H (\mathbf{x}-\mathbf{c})(\mathbf{x}-\mathbf{c})^T\, \mathrm{d}\mu}{\int_H \mathrm{d}\mu}$$

- Solving the algebraic relation between the moments is performed to obtain the relative rotation and translation.

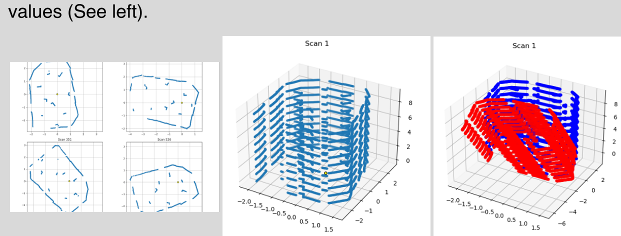
$$\mathrm{c}_2 = \mathbf{R}\mathbf{c}_1 + \mathbf{p}, \qquad \mathbf{\Sigma}_2 = \mathbf{R}\mathbf{\Sigma}_1\mathbf{R}^T$$

The robot was allowed to move in a contained environment with various obstacles to collect LiDAR scans (2D).




Figures 5 and 6: Environmental setup and resulting scan

Resulting 2D LiDAR measurements were initially extrapolated to artificial 3D data. The algorithms were first verified on 3D data with manually rotated point clouds (See below). Then, the same algorithms were run on robot's actual 3D data. The performance of the algorithms and PASTA was evaluated by comparing estimated roto-translations with ground-truth values (See left).




## Conclusion & Future Prospects

Successful validation of the performances of ICP, FPFH, and PASTA was completed through their estimated solutions of 3D data. Results from 3D data align with that obtained from 2D, where PASTA achieves comparable error with faster runtimes when most registration algorithms have high error or longer runtimes, especially with non-optimal initial alignments. Work to further optimize PASTA's performance on 3D point cloud data is in the pipeline. Additional prospects include utilizing PASTA to supervise machine learning point registration algorithms, as well as integration into simultaneous localization and mapping frameworks.

## Acknowledgements & References

I would like to extend my sincere gratitude to Jonathan Bunton, Matteo Marchi, and Professor Paulo Tabuada for their continual support and mentorship throughout the program.
A special thanks to the UCLA Fast Track and SURP programs and their respective staff for facilitating this opportunity.

M. Marchi, J. Bunton, B. Gharesifard and P. Tabuada, "LiDAR Point Cloud Registration with Formal Guarantees," 2022 IEEE 61st Conference on Decision and Control (CDC), Cancun, Mexico, 2022
M. Marchi, J. Bunton, Y. Gas, B. Gharesifard and P. Tabuada, "Sharp Performance Bounds for PASTA," in IEEE Control Systems Letters, vol. 7, pp. 2401-2406, 2023, doi: 10.1109/LCSYS.2023.3285514.
François Pomerleau, Francis Colas and Roland Siegwart, "A Review of Point Cloud Registration
L. Carlone, "Lecture 4: Lie Groups – Visual Navigation for Autonomous Vehicles (VNAV)"

**FACULTY ADVISOR**
Jennifer Wilson

**DAILY LAB SUPERVISOR**
Jennifer Wilson

**DEPARTMENT**
Bioengineering

**BRUCE RUFF**

Computational and Systems Biology
2nd Year, UCLA

## Using PathFX to Study Co-Occurring Diseases

### ABSTRACT

Patients often suffer from multiple co-morbid conditions and may simultaneously take multiple drugs, heightening the risk of side effects. Despite significant patterns in co-morbidity and polypharmacy, patient conditions are treated individually instead of concurrently. An increasingly popular approach for identifying co-occurring diseases involves analyzing the genetic networks linked with these conditions, categorizing them according to shared gene associations. We aimed to understand comorbid diseases using our algorithm, PathFX, which predicts drug-induced phenotypes based on drug binding and downstream proteins at both genetic and network levels.

Leveraging the data derived from PathFX for 4264 diseases, we employed the k-means technique to group diseases based on common drug and gene associations. This process created 87 clusters based on shared drug associations and another 87 clusters based on shared gene associations. The diseases within clusters show moderate overlap overall; within drug-based clusters, the average overlap coefficient among diseases' association data was 0.578, whereas within gene-based clusters, this figure was slightly lower at 0.461. This implies that diseases have a greater overlap of proteins within drug networks compared to the overlap of proteins among the diseases themselves. Furthermore, we conducted a comparison between drug and gene-based clusters, finding that the average highest overlap coefficient associated with each cluster was 0.496. This moderate degree of overlap suggests that drug networks associate to diseases through similar network proteins. Future work could focus on analyzing shared diseases in the electronic health record, which would lead to a better understanding of how drugs might predictively treat co-morbid conditions.

UCLA Samueli School of Engineering
SUMMER UNDERGRADUATE RESEARCH PROGRAM

# Using PathFX to Study Co-Occurring Diseases

Bruce Ruff, Jennifer L. Wilson
Department of Bioengineering — University of California, Los Angeles

## Introduction

**Co-morbid or co-occurring diseases** like diabetes and obesity, asthma and allergies, or depression and anxiety are **conditions** that are often **found together**. The ways these diseases interact are understudied, yet vital for holistic healthcare approaches. It is thought that co-occurring diseases share **connections** in **gene networks**. Our approach **quantifies** the genes that share connections among diseases, along with the genes linked to diseases that co-occur in the **drug networks** produced by PathFX.

## Background & Theory

**PathFX** is a network algorithm that **predicts drug-induced phenotypes**.

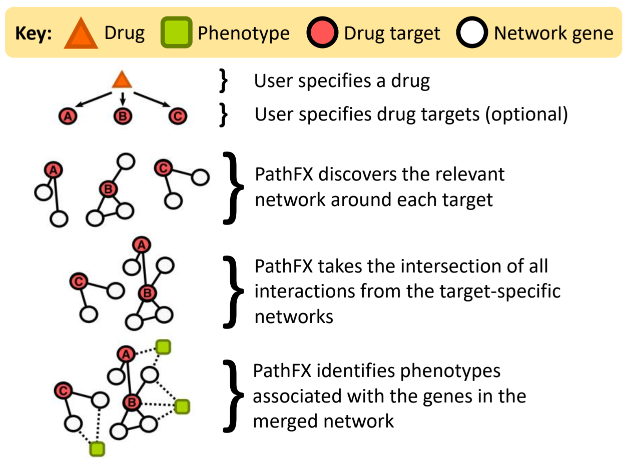



**Figure 1:** PathFX process. For a drug, a gene-interaction network is created to find the most likely phenotypes (Wilson et al., 2019).

***k*-means** is a clustering algorithm that **groups similar data points**. Leveraging PathFX-generated data, we will cluster diseases based on their **drug and gene associations**.

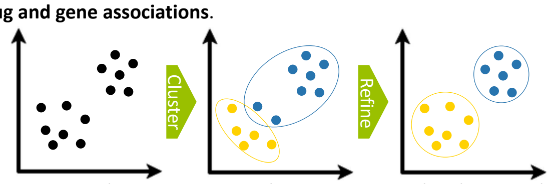
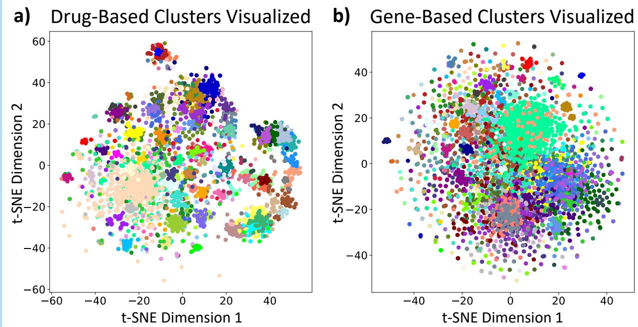



**Figure 2:** *k*-means process. Clusters are created and iteratively refined until data points are optimally grouped.

## Materials & Methods

- **Prepare**
  - Access Hoffman2 computing cluster
  - Run PathFX on drugs in the DrugBank database
- **Link**
  - Construct two dictionaries—one linking diseases to drugs and another linking diseases to genes
- **Cluster**
  - Cluster diseases based on shared drug and gene associations using *k*-means
- **Analyze**
  - Assess intra- and inter- cluster overlap with overlap coefficient, or $|A \cap B| \div \min(|A|, |B|)$
- **Plot**
  - Visualize and compare clusters using SKLearn and NumPy
  - Plot results with Matplotlib

## Results




c) Cluster legend:

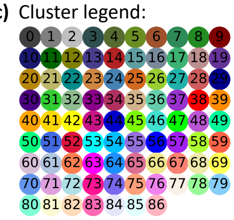



**Figure 3:** Disease clusters visualized via t-Distributed Stochastic Neighbor Embedding (t-SNE) transformation. 4264 disease datapoints are clustered based on drug associations **(a)** and gene associations **(b)**. 87 clusters are assigned distinct colors, as denoted by the legend **(c)**.

Evaluating intra-cluster overlap, disease pairs within drug-based clusters had an average overlap coefficient of 0.578. For gene-based clusters, this average was 0.461.

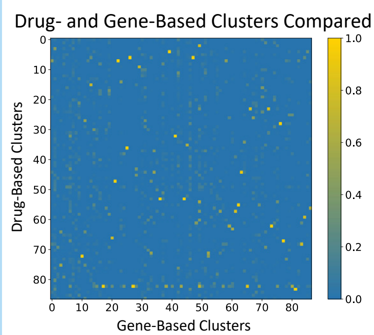



**Figure 4:** Drug- and gene-based clustering discovers distinct disease groupings. Drug clusters are plotted against gene clusters, where overlap is quantified by the overlap coefficient.

Evaluating inter-cluster overlap, the average highest overlap coefficient associated with each cluster was 0.496.

## Conclusions & Future Work

Our work illustrates the ability to **group diseases** based on **shared connections** within their **genetic networks**. Nevertheless, clustering based on drug associations versus gene associations for the same diseases led to different outcomes. Diseases in **drug-based clusters** showed **stronger similarities** than those in **gene-based clusters**. This indicates that drug networks are linked to genes that are common among different diseases, even though the diseases themselves generally have **fewer genes** in common. In the future, data from the **electronic health record** could be leveraged as another method to find co-morbid conditions.

## References & Acknowledgements

Wilson, J. L. et al., (2019). PathFXweb: a web application for identifying drug safety and efficacy phenotypes. Bioinformatics, 35(21), 4504–4506. https://doi.org/10.1093/bioinformatics/btz419

Thank you to Jennifer and the SURP organizers for this opportunity.


## BRYAN LIU

Electrical and Computer Engineering
1st Year, UCLA

**FACULTY ADVISOR**

**Chan Joshi and Sergei Tochitsky**

**DAILY LAB SUPERVISOR**

**Daniel Matteo**

**DEPARTMENT**

**Electrical and Computer Engineering**


# Feasibility Study of 16 µm High-Pressure $Co_2$ Amplifier Optically Pumped at 2.794 µm

## ABSTRACT


There is growing demand for high-power mid-infrared (MIR) laser sources for applications in molecular spectroscopy, atmospheric sensing, and defense. These lasers rely on CO2 for direct amplification. Picosecond 10 µm pulses have been amplified in discharge-pumped systems, requiring expensive high-voltage power supplies at limited repetition rates. Optical pumping of high-pressure CO2 media provides a compact platform to amplify picosecond pulses at 10 µm[1] and nanosecond 16 µm pulses[2]. In this work we investigate the feasibility of amplifying picosecond 16 µm pulses to gigawatt peak powers using a high-pressure optically pumped CO2 medium.

Numerical simulations of ultrashort pulse amplification in an optically pumped CO2-He gas mixture are performed using co2amp[3]. Optical pumping is modeled using a 100 ns pulse at 2.794 µm with 1 J/cm2 fluence, attainable experimentally using an Er:YAG laser. 2 J/cm2 fluence is also considered, approximating a cell pumped longitudinally by two 1 J/cm2 pulses from either side. Optimizations on various system parameters, including amplifier length, CO2/He ratios, and seed wavelength, are performed using µJ-level seed pulses of picosecond duration near 16 µm. Isotopic CO2 gas mixtures are also simulated, allowing amplification of sub-picosecond pulses. In these simulations, we amplify a 10 µJ 16 µm seed pulse to ≥2 mJ of output energy realized in a 10-20 cm long medium with 1 atm of CO2 and ˜10 atm He without considerable increase in pulse duration. Further amplification to ≥10 mJ may be possible if chirped-pulse amplification techniques are adopted.

# Feasibility study of 16 µm high-pressure $CO_2$ amplifier optically pumped at 2.794 µm

Bryan Liu, Daniel Matteo, Sergei Tochitsky, Chan Joshi
Neptune Laboratory | Department of Electrical and Computer Engineering | University of California, Los Angeles

UCLA Samueli School of Engineering

SUMMER UNDERGRADUATE RESEARCH PROGRAM

## Introduction

There is increasing demand for **mid-infrared** (MIR) laser sources for applications in molecular spectroscopy, atmospheric sensing, medicine, and defense. **$CO_2$ gas** is the medium of choice for MIR lasers, able to produce terawatt-level peak power with high efficiency.

## Motivation and Background

- Long wavelength nonlinear optics in gas/semiconductors require high-power sources
- Amplification of ultrashort 10 µm pulses through optical pumping has been studied[1]
- Lasing at **16 µm** pumped by chemical HF laser has been experimentally confirmed[2]
- Upper lasing level pumped at **2.794 µm**, lasing in (021-011) and (020-010) transitions
- Aim to numerically determine feasibility of amplifying an ultrashort 16 µm pulse
  - 2.794 µm pump is experimentally attainable using an Er:YAG laser

## Lasing Transitions

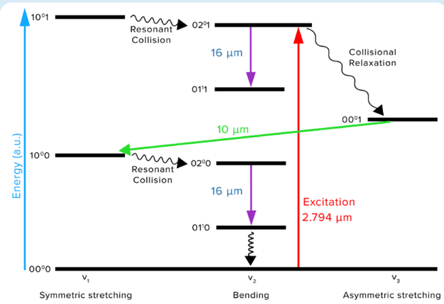
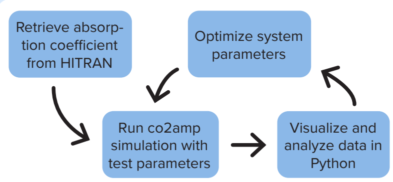



Figure 1. Energy level diagram of the $CO_2$ molecule showing relevant rovibrational levels and transitions for 16 µm lasing

## Materials

- **HITRAN API:** retrieval of absorption coefficients from HITRAN database[3]
- **co2amp software:** simulation of short pulse amplification in a $CO_2$ medium[4]
- **Python:** data visualization and analysis using Numpy and Matplotlib packages

## Methods



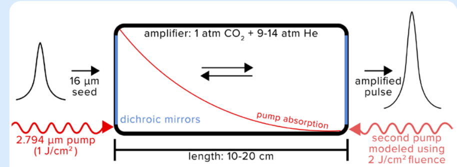




Figure 2. Schematic of optically pumped regenerative amplifier modeled in co2amp

## Results & Discussion

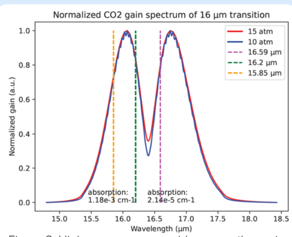



Figure 3. Higher pressures provide a smoother gain spectrum allowing for amplification of ultrashort pulses. Using a 16.59 µm seed also propagates through the atmosphere with minimal energy loss.

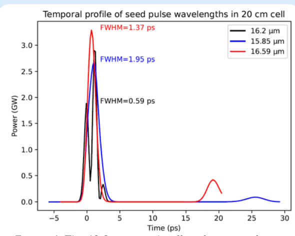



Figure 4. The 16.2 µm seed suffers from significant ultrafast pulse splitting upon amplification, dispersing the energy over multiple pulses. Using a 15.85 µm or 16.59 µm seed largely avoids this issue.

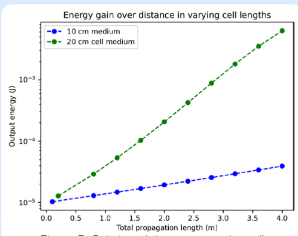



Figure 5. Gain length increases as the cell approaches absorption length. A 20 cm medium absorbs ~95% of energy in the pump pulse and provides a relatively high amount of gain (~1.3%/cm).

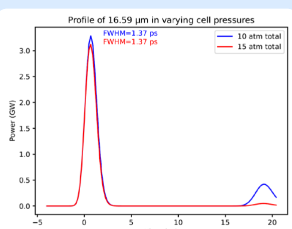



Figure 6. Increasing total cell pressure reduces pulse splitting caused by incomplete overlapping of transition lines with minor losses in gain and ensuring the pulse is as short as possible.

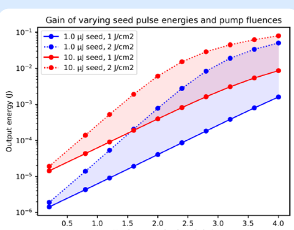



Figure 7. Seed energy is amplified directly by over two orders of magnitude. Even larger gain requiring chirped-pulse amplification (CPA) may be realized by pumping longitudinally from both sides.

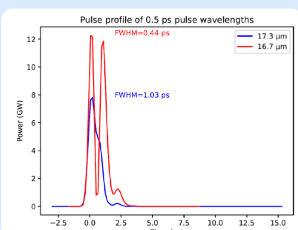



Figure 8. A 60/40 mixture of 626 and 636 $CO_2$ isotopes may be used to amplify a subpicosecond pulse to tens of gigawatts. The seed wavelength is optimized at 17.3 µm to minimize splitting.

## System Parameters

- Pump wavelength: 2.794 µm
- Absorption coeff.: ~0.14 $cm^{-1}$
- Pump fluence: 1-2 $J/cm^2$
- Amplifier length: 10-20 cm
- $CO_2$/He ratio: 1:9-1:14 atm
- Seed wavelength: ~16 µm
- Seed energy: 10 µJ
- Seed duration: 0.5-1 ps FWHM

## References

[1] D. Tovey, S. Ya. Tochitsky, J. J. Pigeon, G. J. Louwrens, M. N. Polyanskiy, I. Ben-Zvi, and C. Joshi, "Multi-atmosphere picosecond CO2 amplifier optically pumped at 4.3 µm," Appl. Opt. 58, 5756-5763 (2019)

[2] M. I. Buchwald, C. R. Jones, H. R. Fetterman, H. R. Schlossberg, "Direct optically pumped multiwavelength CO2 laser," Appl. Phys. Lett. 29, 300-302 (1976)

[3] R. Kochanov, I. Gordon, L.S. Rothman, P. Wcislo, C. Hill, J.S. Wilzewski, "HITRAN Application Programming Interface (HAPI): A comprehensive approach to working with spectroscopic data," Journal of Quantitative Spectroscopy and Radiative Transfer, 177 (2016)

[4] Mikhail N. Polyanskiy, "co2amp: A software program for modeling the dynamics of ultrashort pulses in optical systems with CO2 amplifiers," Appl. Opt. 54, 5136-5142 (2015)

## Conclusions

- Amplification of picosecond & sub-picosecond pulses near 16 µm is possible using a 2.794 µm pump
- A 20 cm, 15 atm (1 atm $CO_2$, 14 atm He) cell is ideal to maximize gain and minimize pulse splitting
- 16.59 µm is minimally absorbed by the atmosphere and is ideal for picosecond pulse amplification
- Amplification from 10 µJ to 1-10 mJ is possible through direct amplification; CPA can amplify to >10 mJ
- A mixture of 626 and 636 $CO_2$ isotopes allows for amplification of 0.5 ps seed pulses

## Acknowledgements

I would like to thank Daniel Matteo and Sergei Tochitsky for introducing me to laser physics and answering the many questions I had over the course of this project. I would also like to thank Dr. Chan Joshi for this opportunity to conduct new research in a field of overlap between engineering and science. I would like to thank my friends and family for their kind words and support this summer. Finally, I would like to thank NSF for providing funding for my project this summer.


FACULTY ADVISOR

**Liang Gao**

DAILY LAB SUPERVISOR

**Qi Cui**

DEPARTMENT

**Bioengineering**

## CHARMAINE TAN

Bioengineering
2nd Year, UCLA


# Optimizing Light-Sheet Microscopy in the Near-Infrared Window for 3D Imaging


## ABSTRACT

Infrared Light Sheet Microscopy has emerged as a powerful imaging technique for researching biological specimens with enhanced penetration depth. In traditional visible light sheet microscopy, imaging depth is limited due to significant light scattering. To address this limitation, we developed a near-infrared open-top light sheet microscopy, utilizing approximately 800 nm excitation, enabling non-invasive, three-dimensional imaging of biological tissues.

To assess the system's capabilities, we imaged various samples of fluorescent beads, zinc agarose, and cross sections from an adult mouse heart. The current microscope system uses a light sheet controlled by a galvanometer to capture multiple images at different focal planes along the z-axis. The current working set-up allows for data collection with an exposure time of 200 ms and within a 400 x 400 µm field of view. Following image acquisition, MATLAB was used to crop the image stack to highlight areas of interest within the samples. Once aligned, the images are stacked into a three-dimensional reconstruction in ImageJ for sample visualization and analysis. Future areas of improvement include decreasing exposure time and maximizing sample field of view to capture morphological nuances in greater detail. Successful integration of light sheet and infrared microscopy in the system will eventually allow for high resolution, 3D imaging of dynamic tissues to research mechanisms of congenital heart disease.

# Optimizing light-sheet microscopy in the near-infrared window for 3D imaging

**Charmaine Tan, Qi Cui, Prof. Liang Gao**

*Department of Bioengineering, University of California, Los Angeles*

UCLA Samueli School of Engineering

SUMMER UNDERGRADUATE RESEARCH PROGRAM

## INTRODUCTION

- Discoveries in modern biological research have been facilitated by advancements in imaging techniques and technology such as **Light Sheet Microscopy (LSM)**. LSM is a non-invasive, high-resolution 3D imaging method that allows observation of biological specimens.
- Traditional visible LSM has limited imaging depth due to significant light scattering.
- We developed a near-infrared open-top light sheet microscope using approximately 800 nm excitation, allowing for 3D imaging of biological tissues at deeper image penetration depth.

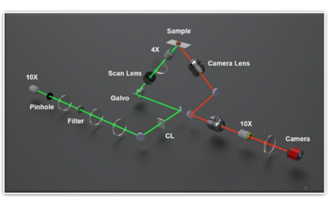
## MATERIALS AND METHODS

### PRINCIPLE

- LSM relies on illuminating a thin section of a sample with a sheet of laser light while capturing emitted light from that plane perpendicularly

### LSM COMPONENTS FOR IMAGING

- **Illumination System**
  A focused laser beam is expanded into a thin light sheet using cylindrical lenses to illuminate a single plane of the sample
- **Detection System**
  Scattered light from the sample is captured using a detection objective placed perpendicularly to the light sheet
- **Galvo Mirror**
  Controls position and angle of scanning beam, allowing light sheet to illuminate planes within the sample



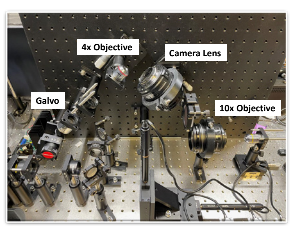




**Figure 1a and b**: Light Sheet Microscope Diagram rendered in SOLIDWORKS and Setup.

### DATA PROCESSING

- **MATLAB**
  Controls galvo mirror to scan light sheet, synchronizes camera with galvo mirror to capture images, and calibrates the LSM system
- **ImageJ**
  Software for 3D image reconstruction

## CONCLUSION AND FUTURE DIRECTIONS

- Our results indicate a functional LSM setup for imaging static phantom and biological samples in the near-infrared window. These sample results are critical for adjusting the setup for future data collection.
- Future areas of improvement include decreasing exposure time from 200ms and maximizing sample field of view to eventually dynamic tissues with markers of congenital heart disease.

## OBJECTIVES

- Build a light sheet microscope for non-invasive deep tissue imaging in the near infrared window
- Assess system capabilities by imaging fluorescent bead, zinc-oxide agarose gel and mouse heart samples
- Test system performance and imaging capabilities in the near-infrared region (NIR)

## RESULTS

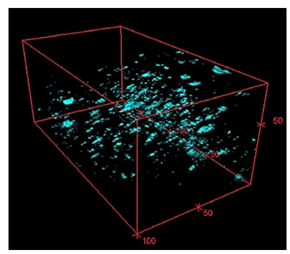
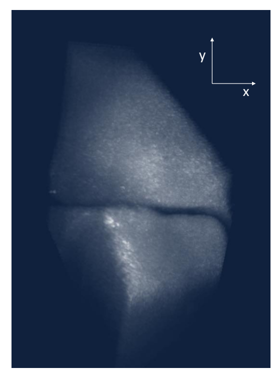
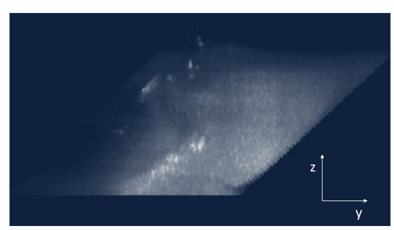


**Figure 2:** Zinc-Oxide Agarose Gel imaged in NIR region, field of view (FOV) 500µm x 500µm x 400µm. 3-Dimensional image reconstructed from image stack with approximately 5µm axial resolution.

**Figure 3a (right) and 3b (bottom).** Cross section from adult mouse heart ventricles imaged in NIR region with 200ms exposure time, FOV 300µm x 400µm x 500µm. Similarly reconstructed from image stack with approximately 6.25µm axial resolution. Sample was irrigated with 1x phosphate buffered saline during imaging.




- Figure 3a depicts x-y plane, Figure 3b depicts y-z plane of sample

**SIGNIFICANCE:**

- Current setup and image reconstruction allowed for visualization of static sample details

## ACKNOWLEDGEMENTS

I would like to extend my gratitude towards my daily lab supervisor, Qi Cui, and my PI, Prof. Liang Gao for their continual support throughout the summer and school year. Additionally, thank you to our collaborators in the Hsiai Lab, SURP, and Samueli Research Scholars for allowing me this opportunity.


**FACULTY ADVISOR**

**Danijela Cabric**

**DAILY LAB SUPERVISOR**

**Enes Krijestorac**

**DEPARTMENT**

**Electrical and Computer Engineering**

**CONNOR STEIGERWALD**

Electrical and Computer Engineering

2nd Year, UCLA


## Co-Secondary Spectrum Sharing in 6 GHz Band with Spatial Prediction

### ABSTRACT


In wireless communications, the 6 GHz frequency band has recently opened for unlicensed secondary users, facilitated by WiFi 6 and cellular 5G technologies. Previously, the 5 GHz band employed listen-before-talk (LBT) to ensure fair channel sharing. LBT entails devices checking for ongoing transmissions before initiating their own data transfer. This study aimed to elevate LBT communication using deep-learning-based radio localization and channel gain spatial prediction. These methods are based on received signal strength measurements from WiFi devices and a 3D map of the environment. The focus was on urban environments, featuring randomly positioned WiFi devices, cellular base stations, and cellular devices. Primary objectives included determining the optimal utilization of the 6 GHz band by cellular base stations amidst WiFi interference and curtailing base station interference on the WiFi through power adaptation.

In our approach, the signal-to-interference-plus-noise ratio (SINR) is calculated at the WiFi and cellular receivers, using the aforementioned predictions. The power of the base stations is adjusted proportionally to minimize interference on the ongoing WiFi transmission. The base stations also decide to operate in the 6 GHz band or opt for a different frequency, based on counteractive interference from the WiFi. Simulations were run, and data was compared with a baseline LBT method, revealing a decrease in interference to Wi-Fi devices and an increase in throughput for cellular devices. The presented work introduces an alternative to traditional LBT methods in the 6 GHz band, with the potential to enhance wireless communication fairness.

UCLA Samueli School of Engineering

SUMMER UNDERGRADUATE RESEARCH PROGRAM

# 6 GHz Co-Secondary coexistence with Spatial Prediction

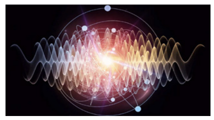
**Connor Steigerwald, Enes Krijestorac, Danijela Cabric**

Electrical and Computer Engineering Department, UCLA

## Introduction

- There has been a recent opening of the 6 GHz frequency band in wireless communications to unlicensed users, which include WiFi 6 and cellular 5G.
- 5 GHz band employed listen-before-talk (LBT) to attempt to achieve fair channel sharing. Where devices check for current transmissions on the medium before transmitting themselves.
- There is an opportunity to improve on LBT to obtain a more fair coexistence where fairness is designated as equal throughput or equal rate to transmit data.


https://community.fs.com/blog/wifi-6-vs-5g.html



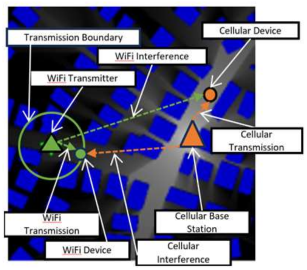
https://www.dudegrows.com/frequencies/

**Objective**: Improve on current LBT method to achieve fair coexistence in the 6 GHz band using neural networks to localize transmitters and predict signal strength

## Materials and Methods

- Python and the Tensorflow library were used to code the neural networks and simulations
- Loc-U-Net RadioLocSeer dataset used for the simulations and the training of the neural networks

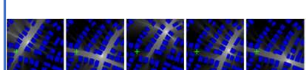
**Simulation Scenario:**
- 256m by 256m random urban environment map
- Five cellular 5G base stations and one WiFi 6 access point are randomly placed in the center of the map
- One cellular base station is picked to transmit to a randomly placed cellular receiver device in its transmission range
- WiFi transmitter sends ongoing transmission to randomly placed WiFi receiver device in its transmission range




**Goal:** Control the transmission of a base station to limit interference onto the ongoing WiFi transmission

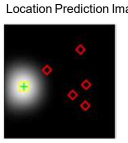
**Assumptions:**
- Know the locations of the base stations and their respective cellular receivers
- Know the range of transmission and power level of WiFi transmitter and base stations
- WiFi transmitter records the received signal strength from the base stations

**Localization Neural Network:**
- Need to know the location of WiFi transmitter
- Loc-U-Net model taken from previous research
- Uses base station predicted radio maps and locations to output a predicted location of the WiFi transmitter
- Data includes set of urban maps with 5 base stations and one transmitter to be localized
- Estimated within 5m of the true location



Example of the map with five base stations and the WiFi transmitter. Green symbol is the transmitter

Location Prediction Image:



Red symbols are base stations, green symbol is true WiFi location, yellow symbol is predicted location, and the grayscale image is the likelihood density heatmap of the predicted location

**LBT Baseline Simulation:**
- Use the same data
- Base station listens to the band, if it hears wifi transmitter, it doesn't transmit
- Collect data on the SINR values and how often the base station transmits

**Approach:**

- Iterate through test data set
- Pick one of the five base stations
- Place cellular device randomly in the base station transmission boundary

↓

- Use simple path loss formulas for the WiFi transmission
- Use the spatial prediction radio maps for the base station transmission

↓

- Within the WiFi transmission boundary, predict the receiver location with the worst case SINR value

↓

- Proportionally adjust the power level of the base station until the SINR at the receiver location is at a certain threshold where the value is seen as effective

↓

- With this new power level, calculate the SINR at the cellular device
- If this SINR for the base station is below the threshold, the base station chooses to not transmit on the 6 GHz band

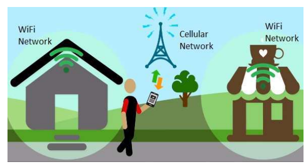
↓

- Collect data on whether the base station does transmit, the SINR value at which it does transmit, and the SINR value for the WiFi transmission

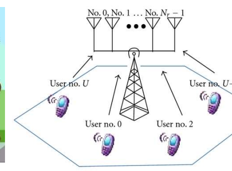
***To get better results would replace the path loss formula used for the WiFi with another spatial prediction trained on ray tracing data**

## Background

- **Wireless Transmissions:** The process of sending and receiving information between devices without physical connections, using radio waves as the medium.
- **Transmitter:** Converts digital data into modulated signals for wireless transmission.
- **Receiver:** Captures and demodulates signals to retrieve original data.
- **Spectrum Sharing:** Efficiently using radio spectrum for multiple wireless technologies to coexist without interference.
- **Frequency Bands:** Allocated spectrum segments for different wireless services.
- **WiFi:** Wireless internet via routers for home, office, and public use.
- **Cellular:** Mobile devices communicate with cell towers for wide coverage and connectivity.
- **Channel Gain:** Signal strength between devices, influenced by distance and obstacles.
- **SINR (Signal-to-Interference-plus-Noise Ratio):** The ratio of the desired signal's power to the combined interference and background noise power. Assesses the quality of the received signal against potential degrading factors.
- **Path Loss:** Signal power reduction over distance due to obstacles and interference sources.



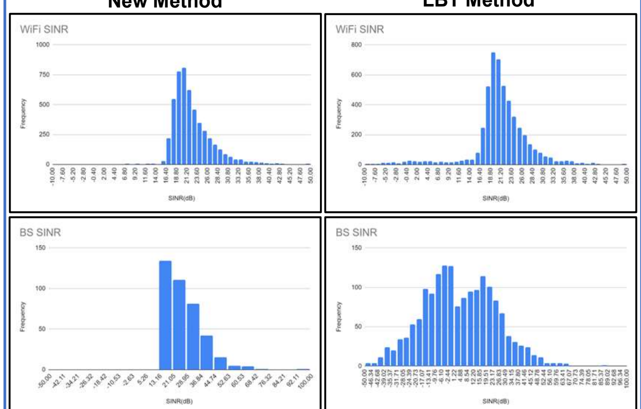

https://www.rfwireless-world.com/Terminology/cellular-data-vs-wifi-data.html




https://www.researchgate.net/figure/Uplink-transmission-in-a-single-cell_fig11_258385079

In this example, devices are transmitting to the cellular base station

## Results




**New Method**

Base Station Transmission Rate: 7.88%
Average WiFi SINR: 22.49 dB
WiFi Collision Rate: 0.84%
Average BS SINR: 27.55 dB
BS Collision Rate: 0%

**LBT Method**

Base Station Transmission Rate: 33.80%
Average WiFi SINR: 20.90 dB
WiFi Collision Rate: 8.06%
Average BS SINR: 4.29 dB
BS Collision Rate: 67.63%

5000 random simulations were run. These results show that the proposed method offers a lower collision rate, where SINR is <15 dB(a designated expected value), for both WiFi and cellular in exchange for less transmissions

## Conclusion

In the results, the proposed method differentiates itself substantially from LBT. In our goal to achieve more fair coexistence, we successfully improved the signal strength of transmissions while simultaneously diminishing the rate of transmission. Both of these variables should increase if fairness is to improve. Thus, the new approach unsuccessfully improves fairness, but offers a potential avenue where signal strengths are much improved opposed to LBT. This research holds potential implications for future endeavors, including the potential to balance signal strength and transmission rate for both 5G and WiFi contexts. Moreover, this methodology could be extended to encompass broader, more intricate real-life scenarios, such as scenarios involving multiple concurrent transmissions. In conclusion, this research presents a stepping-stone to achieving fair co-secondary coexstistence in the 6Ghz band

## Acknowledgments & References

I would like to thank the National Science Foundation and Spectrum X for funding this project through the UCLA Summer Undergraduate Research Program. I would also like to thank the CORES lab and Professor Cabric for giving me this opportunity and for her support and advice throughout the project. Lastly, I would like to thank my lab supervisor Enes who was incredibly helpful and integral to the work on this project.

Yapar, Çağkan, et al. "Real-Time Outdoor Localization Using Radio Maps: A Deep Learning Approach." *IEEE Transactions on Wireless Communications*, Institute of Electrical and Electronics Engineers, Jan. 2023, pp. 1–0, https://doi.org/10.1109/twc.2023.3273202. Accessed 17 Aug. 2023.

Sathya, Vanlin, et al. "Standardization Advances for Cellular and Wi-Fi Coexistence in the Unlicensed 5 and 6 GHz Bands." *GetMobile: Mobile Computing and Communications*, vol. 24, no. 1, Aug. 2020, pp. 5–15, https://doi.org/10.1145/3417084.3417086. Accessed 15 Apr. 2022.

## FACULTY ADVISOR

**Benjamin Williams**

## DAILY LAB SUPERVISOR

**Anthony Kim**

## DEPARTMENT

**Electrical and Computer Engineering**


# DANIEL MCGOVERN

Electrical and Computer Engineering
3rd Year, UCLA


# Laser Feedback Interferometry Using Terahertz Quantum Cascade Vertical External Cavity Laser


## ABSTRACT

The so-called terahertz gap of the electro-magnetic (EM) spectrum has long been characterized by a lack of reliable radiation sources and detectors. The quantum-cascade vertical-external-cavity surface-emitting laser (QC VECSEL) – first demonstrated in Professor Benjamin Williams' lab here at UCLA in 2015 – promises to fill this gap. In laser feedback interferometry (LFI) the external-cavity round-trip time is modulated resulting in a voltage signal from which key parameters of the VECSEL may be obtained. Using the Lang-Kobayashi theoretical model we are able to fit the collected data to the model and estimate one such parameter, the linewidth enhancement factor.

An alternative LFI setup has also been constructed in which the laser is modulated by a low-noise current source and the corresponding voltage signal is measured via a lock-in amplifier. This technique allows for higher SNR (~60 dB) and faster data acquisition. Using this setup and exploiting the bias point dependency of the laser's emission frequency, the LFI signal maps out regions of frequency instability. From this signal we are able to observe at least 843 MHz of frequency tuning over the laser's dynamic range. Unexpectedly, we have observed optical feedback in which the light makes multiple round-trips before being reinjected into the laser cavity – a feature that may be unique to the VECSEL architecture due to its strong susceptibility to feedback. These LFI studies may provide crucial insight for other applications such as optical injection locking.

UCLA Samueli School of Engineering

SUMMER UNDERGRADUATE RESEARCH PROGRAM

# Optical feedback in terahertz quantum-cascade VECSELs investigated through self-mixing interferometry

**Daniel J. McGovern[1], Anthony D. Kim[1], and Benjamin S. Williams[1]**

[1]Department of Electrical and Computer Engineering, University of California Los Angeles (UCLA), Los Angeles, CA 90095

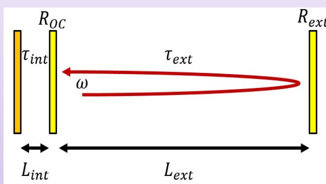
Terahertz and Infrared Photonics Group

## Introduction: Self-mixing

**Laser feedback interferometry (LFI)**, also known as **self-mixing (SM)**, in which light exiting the laser facet ($R_{oc}$) travels an external path ($L_{ext}$), reflects off object ($R_{ext}$) and couples back into laser internal cavity ($L_{int}$).

The phase stimulus, $\phi_s$, represents controlling parameters (laser emission frequency, $\omega_s$, and external cavity travel time, $\tau_{ext}$) affecting the excess phase response, $\phi_{FB}$, of the laser to optical feedback; measurable as changes in laser properties such as frequency, w, terminal voltage, and output power.

$$\phi_{FB} = \omega\tau_{ext}\ ,\ \phi_s = \omega_s\tau_{ext} = \frac{4\pi L_{ext}}{\lambda}$$



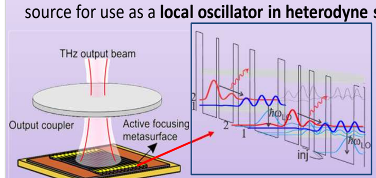

*Derived from the Lang-Kobayashi model the **excess phase equation** governs SM signals:*

$$0 = \phi_{FB} - \phi_s + C\ \sin(\phi_{FB} + \arctan(\alpha))$$

Lang, Roy, and Kohroh Kobayashi. "External optical feedback effects on semiconductor injection laser properties." *IEEE journal of Quantum Electronics* 16.3 (1980): 347-355.

## Introduction: Terahertz (THz) QC-VECSEL

- **Quantum-cascade (QC) vertical-external-cavity surface-emitting lasers (VECSELs)** are composed of a metasurface reflectarray loaded with QC gain material and a high reflectance output coupler.
- THz QC-VECSEL is a single-mode source with the potential to tune in discrete bands over the **1-6 THz range** with milliwatt-level output power.
- Greatly desired for astronomical observations is a frequency stable THz source for use as a **local oscillator in heterodyne spectroscopy**.




- In this work, a **patch-based metasurface** designed at **2.7 THz** is used for lower power consumption and continuous wave operation.

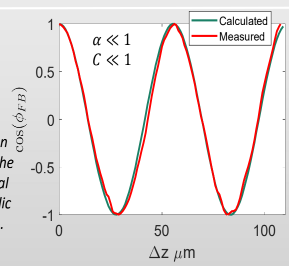
Kim, Anthony D., et al. "Wavelength Scaling of Widely-Tunable Terahertz Quantum-Cascade Metasurface Lasers." IEEE Journal of Microwaves 3.1 (2022): 305-318.

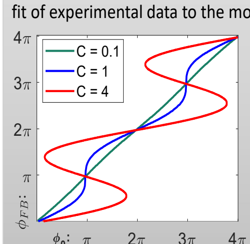
## The Lang-Kobayashi model

**The self-mixing signal is proportional to laser terminal voltage.** $V_{QCL}\big|_{AC} \propto \cos(\phi_{FB})$

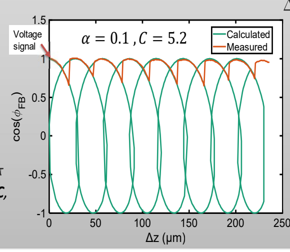
The **linewidth enhancement factor - parameter $\alpha$**; describes coupling between gain and refractive index - may be measured via LFI with a least squares fit of experimental data to the model.

*For a single reflection in $L_{ext}$ the SM signal is periodic with λ/2.*




*For high values of **feedback strength, C**, the relationship between $\varphi_s$ and $\varphi_{FB}$ exhibits path dependence (**hysteresis**).*




$\lambda_s$ = 107 μm
$\omega_s$ = 2.8 THz

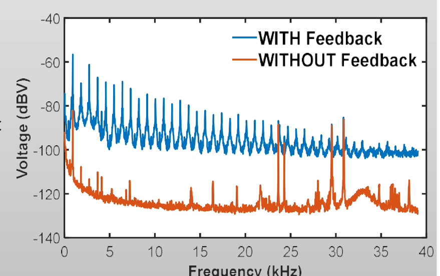
*SM signal λ /4 periodicity is observed. This is consistent with two reflections in the external cavity.*

## The QC-VECSEL under feedback

- Self-mixing has **yet to be demonstrated in the QC-VECSEL**.
- The VECSEL has a highly **gaussian beam shape** resulting in increased coupling efficiency of the re-injected light.
- **High reflectivity of the output coupler** means light may be reflected into the external cavity for multiple round trips.

**Challenges:**

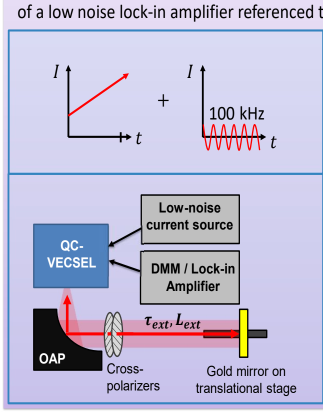
- **Non-monolithic design** of VECSEL cavity causes output coupler vibrations.
- The effect of optical feedback may be to amplify these **mechanical instabilities**.




*FFT of QCL voltage for laser under feedback reveals an increase of ~30 dBV increase of noise power and mechanical resonances separated by ~900 Hz.*

## Self-mixing interferometry via current modulation and lock-in detection

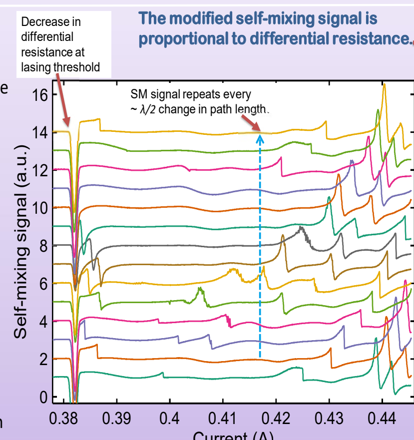
- A current modulation approach, in which the laser bias is swept across the dynamic range, exploits the **bias point dependency of the laser emission frequency** to vary phase stimulus over intervals of 2π.
- Additionally, the bias undergoes **small signal modulation** allowing for use of a low noise lock-in amplifier referenced to the modulation frequency.




- As $\omega_s$ varies across the dynamic range the self-mixing signal periodicity reveals **regions of rapid change in the laser emission frequency** near threshold and the end of the dynamic range; between these two extremes, broad **regions of frequency stability are identifiable**.
- Four periods of oscillation are observed, suggesting ~843 MHz tuning over this bias range.

**The modified self-mixing signal is proportional to differential resistance.**

$$\frac{dV_{QCL}}{dI}\bigg|_{AC} \propto \frac{C\sin(\phi_{FB})}{1 + C\cos(\phi_{FB} + \arctan\alpha)}$$




*SM signal measured via laser bias sweep and current modulation; additionally, $L_{ext}$ is reduced by 5 μm steps; for every ~ λ/2 change in path length the SM signal repeats.*

**Conclusions:**

- Self-mixing interferometry is demonstrated in a QC VECSEL; **experimental data agrees with the Lang-Kobayashi model**.
- The VECSEL is demonstrated to be capable of operating in both low and high regimes of optical feedback.
- An **unexpected λ/4 dependency** is observed in the SM signal.
- Future work may include a study of the LEF as the laser is tuned across the gain bandwidth as well as investigations into laser stability under various feedback conditions.

## Acknowledgements

Funding support for this project is provided by the National Science Foundation and NASA. Daniel McGovern is supported under an NSF REU program supplement to grant 2041165.

## FACULTY ADVISOR
**Yang Zhang**

## DAILY LAB SUPERVISOR
**Xue Wang**

## DEPARTMENT
**Electrical and Computer Engineering**


**DEETSHANA PARTHIPAN**

Electrical and Computer Engineering
1st Year, UCLA


# Watch Your Mouth: Silent Speech Recognition Using Depth Sensing for Smartwatches


## ABSTRACT

Currently, speech-recognition by smartwatches is implemented primarily through audio and voice recognition. This approach is not feasible in situations where users must vocalize privacy sensitive information or in environments with background noise clutter. To solve this, silent speech recognition methods using RGB Cameras, Electrical Conductivity, Pyroelectric effect, and Optical Proximity Sensors have been utilized to model face and lip movement to recognize words. However, these methods are prone to ambient factors such as lighting, background, and skin tone. The goal of this project is to use depth-sensing as a silent speech recognition technique to visualize lip movement and recognize commands.

My goal in the project was to create an interactive website that animates the depth data to visualize lip movement. Depth data was collected in Point cloud format using the True-Depth camera on an iPhone 12 mini and deep learning models such as the YOLO model, and RGB-based visual speech recognition model, AV-HuBERT, were used to correlate the depth maps with phonemes and visemes to recognize commands. Results indicated the system could recognize two pretrained command sets with sizes of 27 and 10 at 82.24% and 85.74% accuracy. I created a website to visualize the lip movement using Django, a python-based backend framework, and a React frontend which utilized the Three.js library to handle animating the Point cloud data. The lip animations for 10 digits and 27 commands(the corpus of data used to train the deep-learning model) from five different speakers can be selected and viewed by the website user. In conclusion, this research shows promise for depth-sensing to enhance the command-recognition abilities of smartwatches.

# Silent Speech Recognition using Wearable Depth Sensing for Smartwatches

Deetshana Parthipan, Xue Wang, Professor Yang Zhang
Human-Centered Computing & Intelligent Sensing Lab
Department of Electrical and Computer Engineering, University of California - Los Angeles

HUMAN-CENTERED COMPUTING & INTELLIGENT SENSING
UCLA Samueli School of Engineering
SUMMER UNDERGRADUATE RESEARCH PROGRAM
FAST TRACK TO SUCCESS
UCLA Electrical and Computer Engineering

## Introduction & Background

**Visemes** – A viseme is any of multiple speech sounds that visually look the same in terms of lip movement

**Phonemes** – Any of the distinct units of sound that differentiate one word from another

- Voice control in smartwatches is primarily implemented through audio recognition which is unideal in situations with background noise clutter or when users must vocalize privacy-sensitive information
- To solve this, silent speech recognition methods using RGB Cameras, Electrical Conductivity, Pyroelectric effect and Optical Proximity Sensors have been researched but are prone to ambient factors such as lighting and skin tone

- Our aim is to utilize a method of silent speech recognition that is unaffected by these external factors
- We identified depth sensing as the most efficient method of silent speech recognition that is unaffected by ambient factors

## Objective

- Use depth sensing to acquire high-fidelity information to recognize user speech
- Create an interactive website that lets use select on certain words/commands and visualize the lip movement for each

## Materials and Methods

- Identified a corpus of data to collect depth data from. We decided on 27 common commands and 10 digits (zero through nine) used frequently in our lives from five different users.

| SCENARIO | COMMANDS |
|---|---|
| Digits Input | Zero, One, Two, Three, Four Five, Six, Seven, Eight, Nine, |
| Music and Podcast | Play, Stop, Resume, Pause, Previous Next, Volume up, Volume Down |
| Smart Home | Turn on, Turn off, Search, Open, Close OK Google, Hey Siri, Alexa |
| System Control | OK, Start, Confirm, Accept, Cancel, Dismiss, Reject |
| Call and Text | Call, Text, Dial, Hang up |

**Figure 1**: Command Corpus

- Collect depth data in Point Cloud format using the True-Depth Camera using an iPhone 12 mini angled at the face.
- We defined 14 visemes (one for silence) and 39 phonemes that have been identified to present the diverse range of sounds in American English in the CMU Pronouncing Dictionary and used a Viseme Phonemes chart (figure #) to help us with Viseme-Phoneme Mapping and Recognition

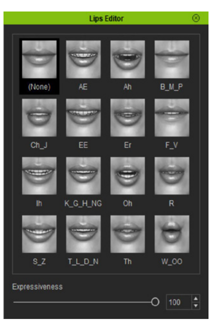



**Figure 2**: Viseme Chart

- To locate lips in depth data we used the pre-trained object detection model YOLOv7
- For viseme detection we used the Point Net model
- We leveraged a pre-trained RGB-based visual speech recognition model, AV-HuBERT, to perform the depth-based silent speech recognition

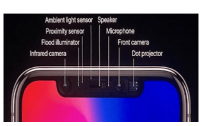



**Figure 3**: True Depth Camera

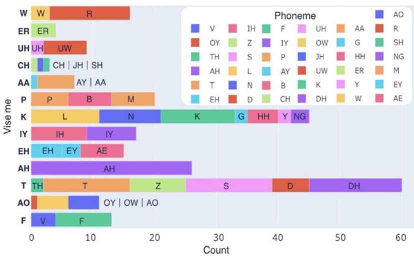



**Figure 4**: Frequency distribution of visemes and phonemes in the Harvard Sentence List2 used to help with Viseme-Phoneme mapping

## Point Cloud Data Collection

**Figure 5**: Data Collection Set-up using iPhone 12 mini

```
[array([[ 0.09688827,  0.12949099, -0.17578125],
       [ 0.09642764,  0.1294145 , -0.17565918],
       [ 0.09585398,  0.12918504, -0.17541504],
       ...,
       [ 0.04583747,  0.11314809, -0.19189453],
       [ 0.04551969,  0.11345423, -0.19238281],
       [ 0.04532117,  0.1140665 , -0.19348145]]),
 array([[ 0.09740333,  0.13017936, -0.17675781],
       [ 0.09694055,  0.13010287, -0.17663574],
       [ 0.09647825,  0.13002639, -0.17651367],
       ...,
       [ 0.04456385,  0.11327054, -0.19213867],
       [ 0.04419433,  0.11345423, -0.19238281],
       [ 0.04382337,  0.11363791, -0.19274902]]),
```

**Figure 6**: A small subsection of the Point Cloud data for one command

For every command there is a separate Point cloud data file (.npz file) where each array represents one frame in the animation of the user speaking the specific word. There are approximately 20 frames for each command. Each array/frame has over 1,000 3D coordinates in the structure (x,y,z) that represent a point that makes up the lip visualization.

## Website Design

- To locate lips in depth data we used the pre-trained object detection model YOLOv7

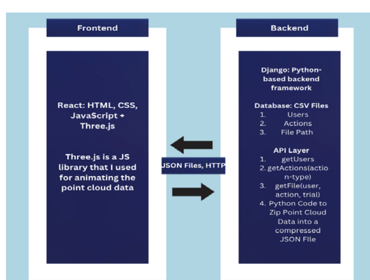



**Figure 7:** System Architecture Diagram

## Three.js Visualizations

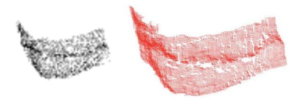


**Figure 8:** Initial Visualizations of Lip Movement

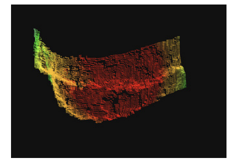


**Figure 9:** Final Visualization of Lip Movement

## Results & Discussion

- The results for user-independent data (fixed position true-depth camera) was 40.44% for commands and 53.50% for digits which was less than expected.
- We also tried obtaining results for user-dependent data (randomly selected one data point from each command and augmented it by rotating the point clouds using a randomly generated rotation matrix). The results for this was 82.24% for commands and 85.74%.
- The website works efficiently. In the beginning the loading time would be longer than desired due to large data transfer. I solved this issue by writing python code to compress the JSON data using gzip and then decompress it in the frontend with JavaScript.

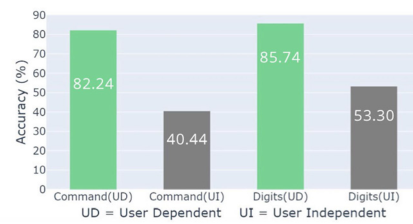



**Figure 10:** Results for User Dependent and User Independent Data

## Conclusions

Results showed that the system could recognize two pretrained command sets with sizes of 27 and 10 at 82.24 % and 85.74% accuracy. These results show potential for further development in the future. However, the accuracy is not sufficient for practical use at the moment. In conclusion, this research shows promise for depth-sensing to enhance the command-recognition abilities of smartwatches.

### Next Steps

- Enhance the website by making the animation movable in the 3D Aspect
- Improve performance with more training data

## Acknowledgements & References

I would like to thank Dr. Yang Zhang, everyone in the HiLab and the UCLA Summer Undergraduate Research Program for their continued support throughout the project. I would also like to thank Xue Wang for her knowledge, guidance, and willingness to help me. Thank you to the National Science Foundation for funding this research. Finally, a special thanks to Will Herrera and all the SURP staff for making this program possible.

[1] "Apple's TrueDepth Camera System." *What Is the TrueDepth Camera System?*, IphoneFAQ, https://www.iphonefaq.org/archives/976228. Accessed 22 Aug. 2023.
[2]https://manual.reallusion.com/Motion_LIVE_Plugin/ENU/Content/iClone_7/Images/Auto_Lip_Options_03.jpg


**FACULTY ADVISOR**
**Greg Pottie**

**DAILY LAB SUPERVISOR**
**Jeffrey Jiang**

**DEPARTMENT**
**Electrical and Computer Engineering**



## EMILY KUCZYNSKI

Electrical and Computer Engineering
2nd Year, UCLA



## KEVIN HONG

Electrical and Computer Engineering
1st Year, UCLA


# Predicting Learning Trajectories with Reinforcement Learning

## ABSTRACT


While intelligent tutoring systems (ITSs) can use information from past students to personalize instruction, each new student is unique—with their own learning style. ITSs should therefore be able to interact with students in a way that maximizes their learning success while respecting their time. By turning to reinforcement learning (RL) algorithms, ITSs can both learn and adapt to new students. Practically, the education problem is partially observed and partial observability typically increases the difficulty of RL significantly; so, we explore what happens when we can use probing interventions to get more information. Gathering information through these interventions reduces the difficulty of final estimation, but it also introduces a cost-benefit decision on how often we want to probe versus help. As a result, our solution seeks to find a balance between probing enough to get accurate estimates and probing so often that it becomes disruptive to the student.

We develop a dynamic, time-series environment to simulate a classroom setting, with student-teacher interventions—including tutoring sessions, lectures, and exams. We evaluate the efficacy of standard RL algorithms under several degrees of partial observability. Our results across Q-learning, Deep Q-learning, XGBoost, and Random Decision Forests (RDFs) are varying but demonstrate effective learning algorithms. In addition, Deep Q-learning, XGBoost, and RDFs are more resistant to changes in partial observability than Q Learning. The models that we develop using these learning algorithms can be used to project when a student should obtain assistance based on their learning trajectory.

# Predicting Learning Trajectories with Reinforcement Learning

UCLA Samueli School of Engineering
SUMMER UNDERGRADUATE RESEARCH PROGRAM

**Kevin Hong, Emily Kuczynski, Jeffrey Jiang, Prof. Gregory Pottie**

Department of Electrical and Computer Engineering, University of California - Los Angeles

## INTRODUCTION

Intelligent Tutoring Systems (ITSs) personalize instruction using past student data. They must interact effectively to enhance learning while respecting the student's time. Reinforcement learning (RL) helps ITSs adapt to new students, but the education problem's partial observability complicates RL. Therefore, we investigate the use of probing interventions to gather additional information, which eases estimation of student learning trajectories but requires balancing the frequency of probing for accurate results without disrupting the student.

## OBJECTIVE

We aim to develop a dynamic, time-series environment to simulate a classroom setting. We will generate and evaluate the efficacy of standard RL algorithms under varying degrees of partial observability.

## PRINCIPLES

Reinforcement Learning is a machine learning paradigm where an agent learns to make decisions by interacting with an environment, receiving feedback in the form of rewards and optimizing its actions to maximize cumulative reward over time [1]. There are three main reinforcement learning algorithms [2].

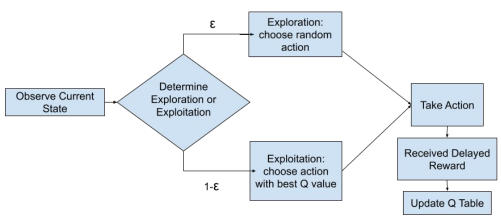



Fig 1: Q-learning Algorithm

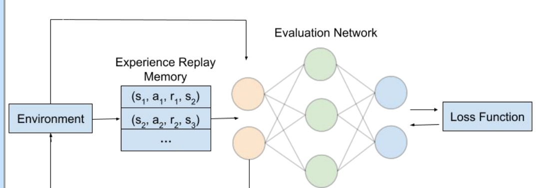



Fig 2: Deep Q-learning Algorithm

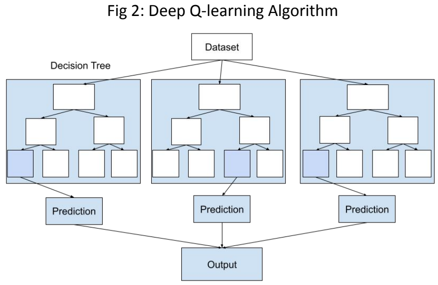



Fig 3: Decision Forest Algorithm

## MATERIALS

**Software Environment:** Python, PyTorch, OpenAI Gym. Scikit Learn
**RL Algorithms:** Q-Learning, Deep Q-Learning, Random Decision Forests (RDFs), XGBoost

## METHODS

In the environment, students engage with multiple courses, each comprising of lectures, a midterm and a final, covering various fundamental concepts. Throughout the courses, the ITS employs various intervention strategies (i.e. tutoring, probing, motivating) depending on individual needs. Students vary in their amount of prior knowledge, level of time commitments to each course and behavioral motivation.

We evaluate the efficacy of the RL algorithms under varying types of partial observability. For some algorithms, we trained the ITS with and without the student's motivation observable. For other algorithms, we included or excluded the lumos intervention as a probe to gain information about the student's expected mastery of the course.

## RESULTS

Fig 4: Q-learning Test Rewards Over Epochs

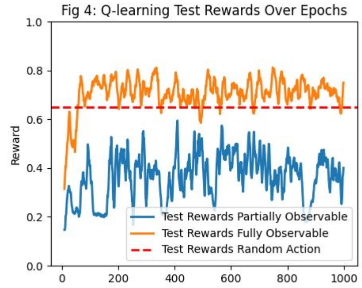



Fig 5: Deep Q-learning Test Rewards Over Epochs

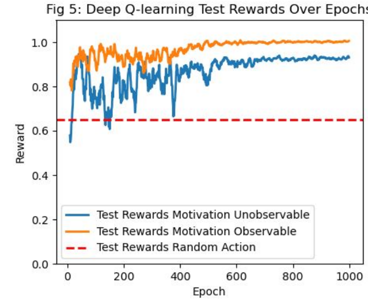



Fig 6: RDF Test Rewards Over Epochs

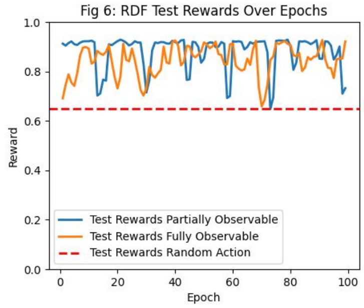
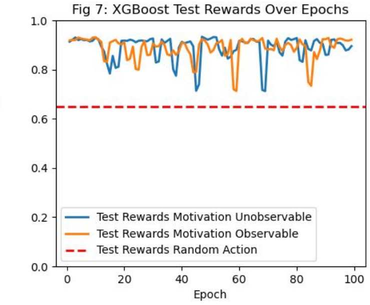



Fig 7: XGBoost Test Rewards Over Epochs




## DISCUSSION & CONCLUSION

Our results across Q-learning, Deep Q-learning, RDFs and XGBoost are varying but demonstrate effective learning algorithms, since the test rewards (a measure of how effective the ITS is at intervening) are above the 0.65 threshold. In addition, Deep Q-learning, XGBoost and RDFs are more resistant to changes in partial observability than Q-Learning: the reward values under the two different levels of observability are much more comparable.

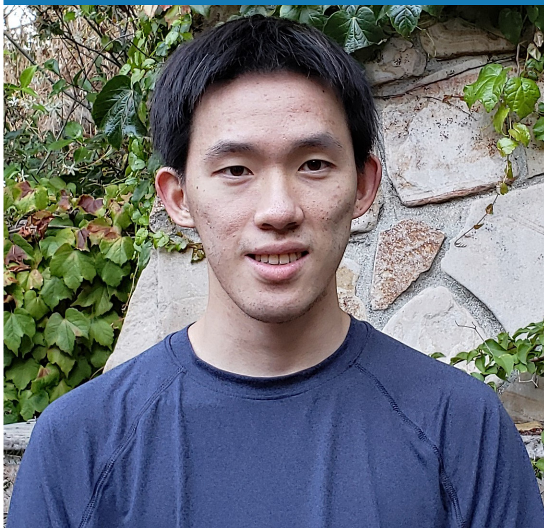
The learning algorithms we employ enable the creation of models that can help predict when a student should seek support according to their learning trajectory.

## REFERENCES

[1] Harmon, Mance & Harmon, Stephanie. (2000). Reinforcement Learning: A Tutorial.
[2] Sutton, R. S., & Barto, A. G. (2018). Reinforcement Learning: An Introduction (2nd ed.). The MIT Press.

## ACKNOWLEDGEMENTS

We would like to thank the Interactive Systems REU for funding our project through the UCLA Summer Undergraduate Research Program. We would like to thank Professor Pottie and our DLS Jeffrey Jiang for their extensive knowledge and support.



## FACULTY ADVISOR
**Baharan Mirzasoleiman**

## DAILY LAB SUPERVISOR
**Yihao Xue**

## DEPARTMENT
**Computer Science**


# ERIC GAN

Computer Science
3rd Year, UCLA


# Which Features are Learned by Contrastive Learning?


## ABSTRACT

Contrastive learning (CL) has emerged as a powerful technique for representation learning, with or without label supervision. However, supervised CL is prone to collapsing representations of subclasses within a class by not capturing all their features, and unsupervised CL may suppress harder class-relevant features by focusing on learning easy class-irrelevant features; both significantly compromise representation quality. Yet, there is no theoretical understanding of class collapse or feature suppression at test time. We provide the first unified theoretically rigorous framework to determine which features are learnt by CL.

Our analysis indicate that, perhaps surprisingly, bias of (stochastic) gradient descent towards finding simpler solutions is a key factor in collapsing subclass representations and suppressing harder class-relevant features. Moreover, we present increasing embedding dimensionality and improving the quality of data augmentations as two theoretically motivated solutions to feature suppression. We also provide the first theoretical explanation for why employing supervised and unsupervised CL together yields higher-quality representations, even when using commonly-used stochastic gradient methods.

# Which Features are Learnt by Contrastive Learning? On the Role of Simplicity Bias in Class Collapse and Feature Suppression

Yihao Xue, Siddharth Joshi, Eric Gan, Pin-Yu Chen, Baharan Mirzasoleiman

UCLA Samueli School of Engineering

SUMMER UNDERGRADUATE RESEARCH PROGRAM

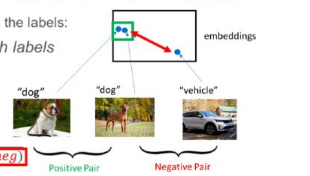
Can we learn better representations?

We need to first understand how and why class collapse and feature suppression happen.

## Introduction

Contrastive learning (CL) has become one of the best representation learning approaches, achieving state-of-the-art performance across various tasks.

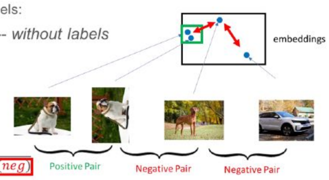
What we can learn if we have all the labels:
Supervised CL (SCL) --- *with labels*




loss function:

$$-\log \frac{\exp \operatorname{sim}(pos)}{\exp \operatorname{sim}(pos) + \sum \exp \operatorname{sim}(neg)}$$

What we can learn without labels:
Unsupervised CL (UCL) --- *without labels*



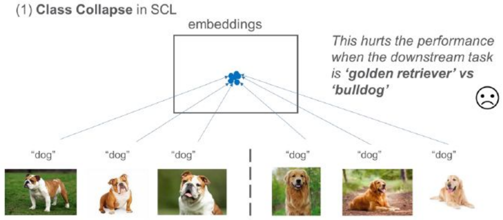

loss function:

$$-\log \frac{\exp \operatorname{sim}(pos)}{\exp \operatorname{sim}(pos) + \sum \exp \operatorname{sim}(neg)}$$

But it may not always achieve satisfactory performance ...

## Two Failure Modes of CL

(1) **Class Collapse** in SCL




(2) **Feature Suppression** in UCL

downstream task: *dog vs car*

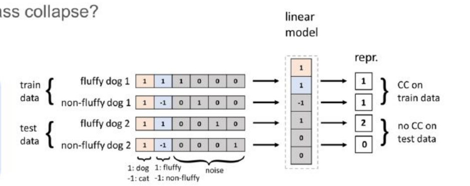
| Features: | dog, moving | dog, still | car, still | car, moving |
|---|---|---|---|---|
| We want the model to learn: | All the features, or at least dog vs. car. | | | |
| When FS happens the mode learns: | moving | still | still | moving |

## Simplicity Bias Contributes to Class Collapse in SCL

Q1: Do all minimizers exhibit class collapse?

A1: No.

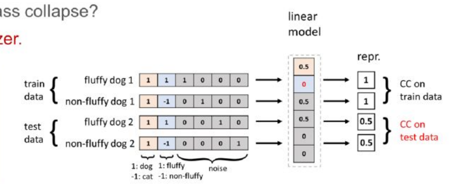
**Theorem (informal)**: ∃ a minimizer of the training loss, s.t. it **learns** the **subclass features** and separates subclasses well on the population.




Q2: What minimizers exhibit class collapse?

A2: The minimum norm minimizer.

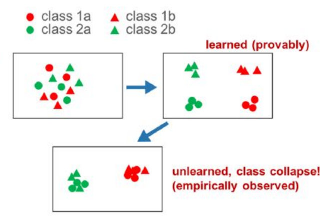
**Theorem (informal)**: The minimum norm minimizer **does not learn** the **subclass features** at all and therefore exhibits class collapse on the population.




Q3: What if we minimize the loss using (S)GD?

A3: Subclasses are **learned** and then **unlearned**.

**Theorem (informal)**: In GD, there exists an epoch where **subclass features** are **learned** and subclasses are well separated in the embedding space.




Q4: What causes class collapse in (S)GD?

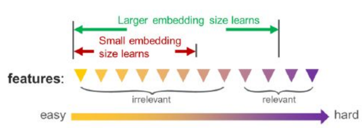
Conjecture: the optimization algorithm's bias toward simple (e.g., min norm) solutions.

## Understanding Feature Suppression in UCL

**Many factors** can contribute to feature suppression.

1. Embedding size:

**Theorem (informal)**: With (1) easy-to-learn task-irrelevant features and (2) insufficient embedding size, the **min norm minimizer** exhibits feature suppression.



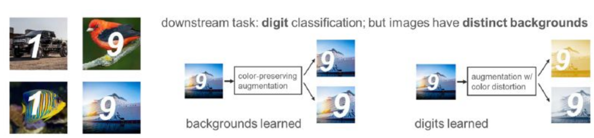

This suggests **increasing embedding size** as a solution

E.g., larger embedding size leads to better downstream performance on CIFAR10-RandBit

| Embedding size | Downstream accuracy |
|---|---|
| 4 | 86.73 |
| 64 | 96.82 |
| 128 | 97.65 |

2. Data augmentation:

**Theorem (informal)**: With (1) highly diverse irrelevant features and (2) imperfect data augmentation, the **min norm minimizer** exhibits feature suppression, even with arbitrarily large embedding size.




## Joint Loss Can Mitigate Both Issues

We provide the *first theoretical justification* for the joint loss.

**Joint loss** = $\beta$ * Supervised CL loss (prioritize class features) + $(1-\beta)$ * Unsupervised CL loss (encourage learning of other features)

**Theorem (informal)**: The joint loss can avoid both class collapse and feature suppression.

E.g., joint loss leads to better class and subclass accuracies on CIFAR100-RandBit

| Loss | Subclass acc | Class acc |
|---|---|---|
| SCL | 28.1 | 61.1 |
| UCL | 34.1 | 52.3 |
| Joint | 35.7 | 63.9 |

Acknowledgements

I would like to thank NSF and other funding sources for funding the UCLA Summer Undergraduate Research Program. I would like to thank my collaborators and Yihao Xue, Siddharth Joshi, Pin-Yu Chen, and Professor Baharan Mirzasoleiman for their resources, knowledge, insight, and support. I would like to thank the SURP staff for their guidance.

This work is based on a paper in ICML 2023 of the same name.


## EUGENIA CHO

Electrical and Computer Engineering
1st Year, UCLA


**FACULTY ADVISOR**

**Christina Fragouli**

**DAILY LAB SUPERVISOR**

**Merve Karakas and Xinlin Li**

**DEPARTMENT**

**Electrical and Computer Engineering**

# Enhancing Online Learning-Based Sensor Selection in Vehicle Tracking Using Maximum a Posteriori Detector (MAP) with Elimination


## ABSTRACT

Vehicle tracking technology holds diverse applications across various industries worldwide. Most commonly, vehicle tracking technology uses a combination of sensor nodes, which capture relevant location data, and computational units, which process the sensor data, to execute vehicle tracking. The underlying assumption in the context of vehicle tracking is the availability of appropriate sensor nodes. While the simplest approach to fulfill this assumption is to activate all sensors at all times, doing this proves impractical due to cost and efficiency concerns, especially if the area of interest is expansive. The lab thus focuses on exploring algorithms that decide when to activate specific sensors and has shown that the Maximum A Posteriori with Arm Elimination (MAP_e) algorithm outperforms other multi-armed bandit algorithms in terms of sample cost, accuracy, and latency.

Our work involved simulating a vehicle tracking scenario using Python. The simulation depicted a grid representing the area of interest, with a combination of available and blocked (unavailable) cells, allowing the vehicle to move in specific directions. We provided extensive simulation results where we explored a number of scenarios and algorithms, including our proposed improvements to the algorithm as well as the consideration of error propagation while tracking vehicles.

UCLA Samueli School of Engineering

SUMMER UNDERGRADUATE RESEARCH PROGRAM

# Enhancing Online Learning-Based Sensor Selection in Vehicle Tracking Using Maximum a Posteriori Detector (MAP) with Elimination

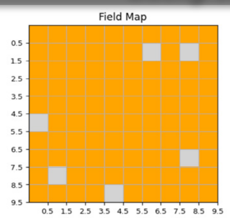
**Eugenia Cho, Merve Karakas, Xinlin Li, Christina Fragouli**

Department of Electrical and Computer Engineering, University of California, Los Angeles

## Introduction

In applications of vehicle tracking, the assumption that necessary sensors for tracking are already available often overlooks the critical step of selecting the relevant sensors. Particularly when the area of interest is vast, inefficient **sensor selection** leads to suboptimal outcomes across various metrics including **latency and communication cost.** Thus, in order to select appropriate sensors based on the most accurate prediction of the vehicle's trajectory, we incorporate **online learning**, where a model learns data as it becomes available, and propose the **Maximum a Posteriori detection with arm elimination (MAP_e)** algorithm, which adapts ideas from a subset of **multi-armed bandit** algorithms called **best-arm identification.**

## Background and Objective




**Sensors** are deployed in all orange cells and are naturally partitioned into groups. We will investigate which groups to activate at which time.

Fig. 1: abstract depiction of a field map

### Multi-Armed Bandit (MAB) Problem

An agent is faced with a set of options, or "arms", where each arm provides a reward based on some probability distribution. The agent makes a series of choices under uncertainty while facing an exploration-exploitation trade-off in order to maximize the cumulative reward.

### Best-Arm Identification (BAI)

BAI is a specific scenario within the MAB framework where the primary objective is to identify the arm with the highest expected reward as opposed to maximizing cumulative reward

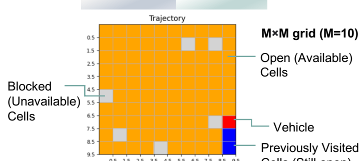
**Objective**: Further improve the MAP algorithm under complex scenarios such as diagonal movements, error propagation, and expanded action sets, where we expect elimination to help with overall metrics.

## Materials and Methods

Step 1 Set Up

Experiments were run locally on Visual Studio Code using Python

Step 2 Model




Step 3 Test

Compare the **accuracy** (percentage of correct tracking), **latency** (number of time slots needed to decide on sensors) and **sample cost** (total number of samples collected) under different scenarios

### Maximum a Posteriori with Arm Elimination (MAP_e)

Collect data from nearby candidate sensors and factor in the prior probability distribution of possible arms (locations). The MAP detector selects the arm with the highest posterior probability to activate the sensors, and MAP_e eliminates possible arms that fall below a predetermined threshold to reduce latency and sample cost.

$$\begin{aligned} l_{\text{MAP}}(t) &= \arg\max_{l \in L_t} \mathbb{P}[l_t = l \mid \{a_i\}_{i=1}^n, \{r^{(i)}\}_{i=1}^n] \\ &= \arg\max_{l \in L_t} \mathbb{P}[\{r^{(i)}\}_{i=1}^n \mid l_t = l, \{a_i\}_{i=1}^n]\,\mathbb{P}[l_t = l] \\ &= \arg\max_{l \in L_t} \mathbb{P}[l_t = l] \prod_{i=1}^n \mathbb{P}[r^{(i)} \mid l_t = l, a_i], \end{aligned}$$

Fig. 2: how MAP detection decides on the location of a tracked vehicle

## Results

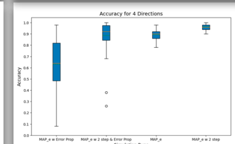



Fig. 3: accuracy comparison between MAP_e and MAP_e with second step

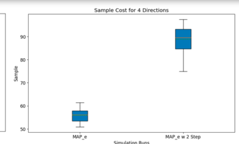



Fig. 4: sample cost comparison between MAP_e and MAP_e with second step

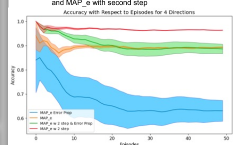



Fig. 5: visualization of accuracy variance across each time slot

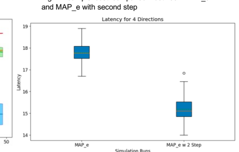



Fig. 6: latency comparison between one step and two step sample size

***Comparison between considering 4 immediate next steps v. considering subsequent next steps***

**Fig. 3-6** depict the performance of MAP_e across various metrics when factoring in second next step into mu calculation, where elimination provides an advantage overall

***Comparison between 4 possible vehicle movement directions v. 8 possible movement directions***

**Fig. 7-10** depict the performance of MAP v. MAP_e and 4 v. 8 directions in terms of latency and sample cost, where expanding the action set requires more sampling to reach comparable accuracy

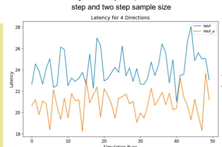



Fig. 7: difference in latency between MAP and MAP_e with 4 possible directions

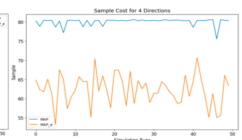



Fig. 8: difference in number of samples required for MAP and MAP_e with 4 possible directions

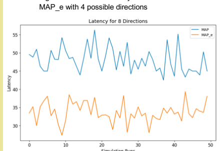



Fig. 9: difference in latency between MAP and MAP_e with 8 possible directions

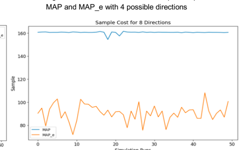



Fig. 10: difference in number of samples required for MAP and MAP_e with 8 possible directions

## Conclusion

Expanding sampling to the **second set of possible next steps** enhances the algorithm's vehicle trajectory prediction accuracy, particularly in **error propagation** scenarios where the true vehicle location is not known at each time slot. MAP_e consistently demonstrates higher accuracy and reduced latency, although the sample cost is nearly doubled as it samples from twice as many cells.

Enabling **diagonal movements** requires more samples compared to the 4-directional approach (left, right, up, down). Despite increased overall latency with diagonal movements, MAP_e outperforms regular MAP in terms of reduced latency and sample cost, solidifying the robustness of our proposed algorithm with arm elimination.

## Future Work

Multiple Vehicles

Consider interference and probability distributions when multiple vehicles share a grid

Recalculate Priors

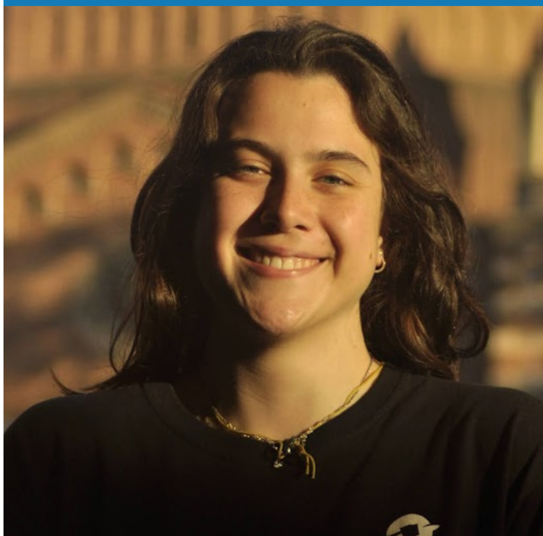
Update reward function and modify how prior probability distribution is calculated

## References

X. Li, M. Karakas, O. Hanna, M. Kiamari, J. Coleman, C. Fragouli, B. Krishnamachari, G. Verma, "Online Allocation of Sensing and Computation," under preparation, 2023

## Acknowledgements

I would like to thank Professor Christina Fragouli and the ARNI Lab, especially Merve Karakas and Xinlin Li, for their guidance this summer. I would also like to thank the Summer Undergraduate Research Program (SURP) for organizing this opportunity and NSF for funding my research through grant 2007714.




## GABRIELLA MUNN

Electrical and Computer Engineering
1st Year, UCLA



## MARCUS SIVAYAVIROJNA

Electrical and Computer Engineering
1st Year, UCLA


**FACULTY ADVISOR**

**Ethan Wang**

**DAILY LAB SUPERVISOR**

**Lap Yeung**

**DEPARTMENT**

**Electrical and Computer Engineering**

# Design of Ultra-Sensitive, Broadband Very Low Frequency Receivers with Equal Potential Coupled Coils

## ABSTRACT


Highly sensitive very low frequency sensing coils are a fundamental component of many radio frequency receiver systems, with practical applications ranging from ultra-low field MRI to underwater signal detection. However, the traditional cascaded coil design has begun to reach its limitation in regard to sensitivity. Due to ohmic resistance, the voltage drops throughout a coil, introducing parasitic capacitance between each turn. Parasitic capacitance lowers the inductance of the coil, directly limiting coil sensitivity. This project aims to test a novel design known as the equal-potential coupled (EPC) coil. By splitting a coil into segments and applying an equal voltage to each segment, all segments are equipotential. This reduces parasitic capacitance, improving coil sensitivity. Using Ansys HFSS and CST Studio, we iteratively modeled and analyzed various coil designs. Maximizing the Q factor leads to improved coil sensitivity, and was therefore the focus of our coil comparisons.

Initially, we hypothesized that the EPC coil's Q factor would be approximately 4x that of the cascaded coil. The simulation results were inconclusive, with the plotted Q factors of the cascaded coil, parallel coil, and EPC coil being very similar. A small decrease in the inductance suggests that proximity effect was present, but very weak due to high spacing between coil turns. Future work will revolve around modeling more compact coils in order to observe the full impact of capacitance on the inductance of different coil geometries.

SUMMER UNDERGRADUATE RESEARCH PROGRAM

# Design of Ultra-Sensitive, Broadband Very Low Frequency Receivers with Equal Potential Coupled Coils

Marcus Sivayavirojna, Gabriella Munn, Prof. Yuanxun Ethan Wang, Dr. Lap Yeung
Department of Electrical and Computer Engineering, University of California, Los Angeles

## Introduction

Certain applications, including low frequency MRIs [1], require highly sensitive receivers. Coils are suitable for this, as their geometry allows them to achieve high sensitivity. The coil sensitivity is a measure of the minimum signal that can be detected, and it can be calculated using the following equation [2]:

$$\delta B = \sqrt{\frac{4\mu_0 kT}{\mu_r \omega v Q_{ind}}} \quad \text{where} \quad Q_{ind} = \frac{\omega L}{R}$$

A coil's resistance results in voltage to drop throughout the coil. The voltage different between the coil's turns causes the coil to act like a capacitor, resulting in the current flowing through the coil to be out of phase. This reduces the inductance and $Q_{ind}$, making the coil less sensitive.

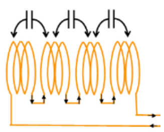


Figure 1. Cascading Coil

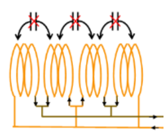


Figure 2. Equal-Potential Coupled Coil

A novel equal potential coupled coil design has been proposed to minimize the capacitance. By parallelizing the current, the parasitic capacitance between the coil subsets can be removed, resulting in a higher coil sensitivity. Our goal was to test this experimental design by simulating a cascading coil, parallel stranded coil, and equal potential coupled coil.

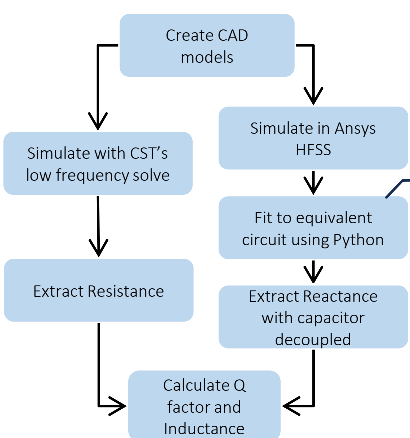
## Materials

**CST Studio**
EM simulation software

**Ansys HFSS**
High Frequency EM simulation software

**Python**
Programming Language

## Method



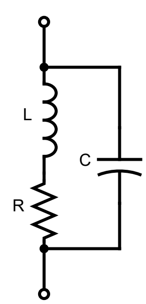



Figure 3. Equivalent circuit for coil


Figure 3: Modelled Cascading Coil

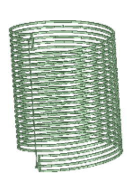


Figure 4: Modelled Parallel Coil


Figure 5: Modelled Equal Potential Coupled Coil

## Results

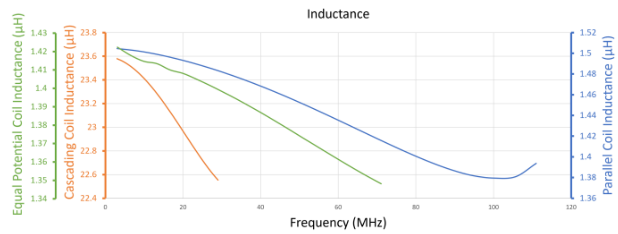


Figure 6. Inductance calculated from imaginary impedance with capacitor decoupled

As the frequency increases, the capacitance begins to impact the inductance. The parallel coil and equal potential coupled coil have similar initial inductances that drop off at similar rates. The cascaded coil's inductance is larger with a much steeper drop.

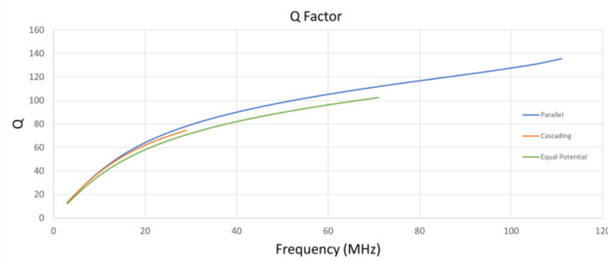


Figure 7. Q factor plot after decoupling the capacitor from circuit

The cascading coil, equal potential coupled coil, and parallel coil resonate at 30, 70, and 110 MHz respectively. The Q factors do not vary drastically between each coil geometry.

## Conclusion

The results align with the generally accepted theory surrounding the impact of parasitic capacitance on coil inductance. The inductance of all three simulated coils decrease as frequency increases, highlighting the negative effect of parasitic capacitance on coil inductance.

It was originally hypothesized that the equal potential coupled coil would perform much better than the cascaded coil design, with an expected 4x increase in Q factor. Yet in the simulated results, the Q factor did not change significantly between all three simulated coil. It could be concluded that the parasitic capacitance's influence on the Q factor is minimal.

However, coils are typically tightly wound with significantly less spacing between the turns. This would increase the parasitic capacitance and its impact on the inductance, potentially changing the results. Future simulations will need to be done with more tightly wound coils in order to obtain more conclusive results.

## References

[1] Savukov, I., Kim, Y. J., & Newman, S. (2022). High-resolution ultra-low field magnetic resonance imaging with a high-sensitivity sensing coil. *Journal of applied physics*, *132*(17), 174503. https://doi.org/10.1063/5.0123692
[2] Wang, Yuanxun E. (n.d.). *Inductive sensor sensitivity derivation and comparison.* [Unpublished manuscript]. Samueli Electrical and Computer Engineering Department, University of California, Los Angeles

## Acknowledgements

We would like to thank everyone in the Digital Microwave Lab, with special thanks to Dr. Lap Yeung and Professor Yuanxun Ethan Wang for their support and guidance throughout this project. We would like to thank the National Science Foundation for funding our project through the UCLA Summer Undergraduate Research Program.


**FACULTY ADVISOR**

**Aaswath Raman**

**DAILY LAB SUPERVISOR**

**David Abraham**

**DEPARTMENT**

**Electrical and Computer Engineering**

**GWYNETH LEMASTER**

Electrical and Computer Engineering
2nd Year, UCLA


## Physics-Enhanced Neural Networks for Nanophotonic Structure Design


### ABSTRACT

A central challenge in modern optics and materials research lies in identifying and designing the optimal nanostructure that can deliver a desired set of optical properties. Numerical device simulations traditionally harness fundamental physical laws (e.g., Maxwell's Equations) to simulate such nanophotonic structures and remain the standard in the field. However, these methods are computationally expensive, making optimization challenging. Alternatively, recent work has proposed using a neural network that is trained to predict the optical response of a nanophotonic structure, acting as a surrogate simulation method that is much quicker and less computationally expensive. Conventional neural network training methods employ the root-mean-square error loss function, which is a direct measure of discrepancy between neural network predictions and ground truth. This function, while straightforward and robust, lacks a deeper insight into the intricacies of the physics involved.

Moreover, the trained neural networks do not generalize well and are unable to perform simulations beyond the domain of the training data set. By integrating physics-informed metrics into the loss function, we aim to deepen the neural network's understanding of the underlying physics, enhancing the model's predictive capability and efficiency, and ability to generalize. We do this by explicitly incorporating the absorption peak wavelength into the loss function, where the loss is proportional to the difference in the peak-absorption position. This modified approach is designed to accelerate training convergence, reduce the required epochs, and facilitate more efficient learning even with limited datasets. Finally, it is a well-known phenomenon that neural-networks fail to predict the response of inputs outside of the range of inputs in the training set. A physics-informed loss function may enhance a neural-networks ability to generalize beyond the range of the training-set.

# Physics-Enhanced Neural Networks for Nanophotonic Structure Design

Gwyneth Lemaster, David Abraham, Christopher Yeung, Aaswath Raman
*Raman Laboratory, Department of Materials Science & Engineering, University of California, Los Angeles*

UCLA Samueli School of Engineering
SUMMER UNDERGRADUATE RESEARCH PROGRAM

## Introduction

In recent years, machine learning techniques have emerged as alternate strategies for both forward and inverse design (determining the optimal parameters or configurations for a system based on desired output or performance criteria) of nanophotonic structures. As a relatively new tool for photonic inverse design, deep learning (DL) has shown promising results in some applications and by certain measures. In deep learning, convolutional neural networks (CNNs), are a type of neural network most applied to processing grid-like data such as images. CNNs enhance a model's ability to conduct pattern recognition as they offer high predictive power and more efficient training.

In this project, physics-informed metrics and deep learning-based inverse design methods are combined to enhance the performance of conventional CNN-based inverse designs for nanophotonic structures. A custom loss function is used that incorporates the absorption peak wavelength, thereby improving the model's ability to generalize and predict optical properties more efficiently.

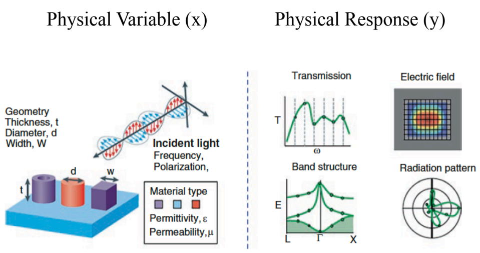



*Figure 1. Design parameters (left) and optical responses (right). Computing the responses y of a device described by given parameters x is known as the forward problem, whereas retrieval of x from y is the inverse problem.*

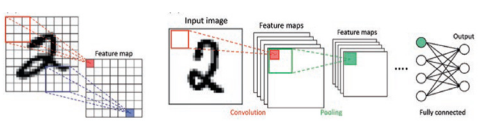



*Figure 2. Typical CNN architecture.*

## Objective

The objective is to apply "inverse design" methods, where in response to target absorption spectra, the network can identify an effective metasurface in terms of its class, materials properties, and overall shape to understand how nano- or micro-scale structures and material properties can be optimally configured to attain specific functionalities.

## References

[1] Yao, K., & Zheng, Y. (2023). Deep-Learning-Assisted Inverse Design in Nanophotonics. In Nanophotonics and Machine Learning: Concepts, fundamentals and applications. essay, Springer Nature Switzerland AG.
[2] Yeung, C. *et al.* (2020) 'Elucidating the behavior of nanophotonic structures through explainable machine learning algorithms', *ACS Photonics*, 7(8), pp. 2309–2318. doi:10.1021/acsphotonics.0c01067

## Materials and Methods

**Python** programming language and **PyTorch, Tensorflow, Scikit Learn** libraries.
**Spyder** is the primary software application used for simulation and modeling tasks.
**Lumerical,** a photonic design and analysis software tool, was utilized to perform three-dimensional finite-difference time-domain (FDTD) simulations of unique structures.

**1.** Facilitate data collection by converting 3D metal−dielectric−metal metamaterials into 2D representations for image-based machine learning.

**2.** Train multiple CNN architectures to predict the electromagnetic response of input images.

**3.** Determine optimum hyperparameters and confirm model performance of saved models using validation and test sets.

## Results and Discussion

In the training of the CNN, the custom loss function (CLF) demonstrated enhanced performance in the inverse design task compared to the mean-square-error loss function (MSE). Specifically, the prediction accuracy improved by approximately 35.85% for the in-range test set and 7% for the beyond-range test set. This improvement suggests that the CLF may capture certain features or relationships in the data that the traditional MSE neglects. As for the beyond-range test set, the improved performance of the CLF could be attributed to its capability to filter out noise, offering a more refined model prediction rather than generalization.

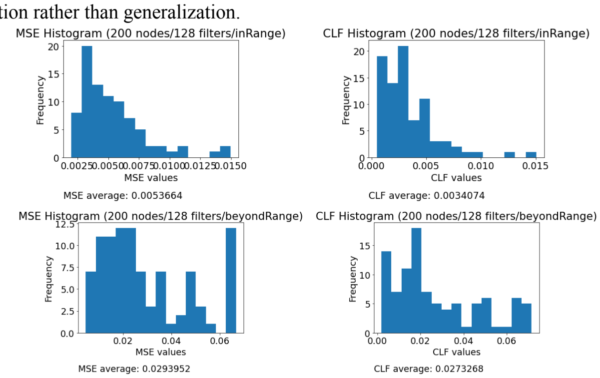



*Figure 3. MSE histograms (left) and CLF histograms (right).*

## Conclusion and Future Work

These findings further demonstrate that when the network can identify the underlying relationships between structure, material, metasurface class, and optical response, there is potential for new yet accurate solutions that extend beyond the known designs.

In the future, investigations of other fundamental optical properties of materials (e.g., real and imaginary refractive indices, magnetic permeability, etc.) can be incorporated into the model to achieve a more generalized inverse design framework.

## Acknowledgements

Thank you to the National Science Foundation for funding this project through the UCLA Summer Undergraduate Research Program. Additionally, professor Aaswath Raman for his resources, knowledge, and support, as well as David Abraham and Will Herrera for their research guidance.


**FACULTY ADVISOR**

**Tyler Clites**

**DAILY LAB SUPERVISOR**

**Sachi Bansal**

**DEPARTMENT**

**Mechanical and Aerospace Engineering**

## HEATHER GRIFFITHS

Materials Engineering
3rd Year, UCLA


# Design of an Anthropomorphic Thumb Tip for an Osseointegrated Prosthesis


## ABSTRACT

Opposition is a unique function attributed to the shape and position of the thumb, and is an essential component of hand function that thumb prostheses must be able to perform. When designing a prosthetic device, it's important that the appearance of the prosthetic thumb, including shape and texture, mimic a biological thumb to help support function and to allow for the user to perceive the device as part of the body. This process, called embodiment, is supported by creating anthropomorphic, or human-like devices. This research project is built off of the development of a fully-implantable thumb prosthesis that will be covered by the patient's skin in order to help restore the ability for the patient to sense texture and temperature, allowing for embodiment.

The goal of this research project was to test different thumb tip shapes for the fully-implantable thumb prosthesis to see which shape is the most anthropomorphic. AutoCad was used in order to create and model various thumb tip shapes based on radiographic data of the thumb, then the parts were 3D printed. Next, cadaver forearm skin was placed on top of the parts to assess how the parts looked underneath human skin, and to see which looked the most anthropomorphic. There were a total of eight thumb tips printed and used in the cadaver forearm dissection.

UCLA Samueli School of Engineering

SUMMER UNDERGRADUATE RESEARCH PROGRAM

# Design of an Anthropomorphic Thumb Tip for an Osseointegrated Prosthesis

Heather Griffiths, Sachi Bansal, Tyler R. Clites

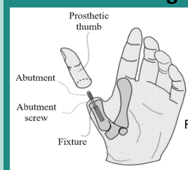
Department of Materials Science and Engineering University of California — Los Angeles

Department of Mechanical and Aerospace Engineering, University of California — Los Angeles

## Introduction

The thumb is a very important part of the hand since it opposes the other fingers on the hand, allowing us to pinch and grasp objects. However, traumatic finger amputations, including the thumb, make up more than 90% of amputations in America [1]. Therefore, it's important that researchers creating thumb replacements implement solutions that allow patients to replicate these motions.

### Osseointegration




Osseointegrated prostheses: Devices directly implanted into the amputee's bone

Osseoperception: Restores some of the tactile sensations in osseointegrated prosthesis

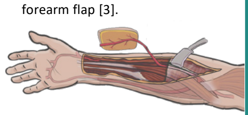
Figure 1. An osseointegrated prosthesis and its assembly [2].

### Skin Flap Transfer

Prelamination: The prosthetic device is placed underneath the flap in Fig. 2 for ~ 4-6 weeks before the transfer for the skin to adhere to the prosthesis [4].

- Reduces *extrusion*, or repetitive motion that pushes the prosthetic device through skin

Figure 2. A dissection depicting the removal of the radial forearm flap [3].



Skin flap transfer: The prosthetic device is removed and placed onto the thumb with the skin flap around it.

- Increases stability, and sensory and tactile feedback in the device

### Shape and Function

A prosthetic device's shape and function are intertwined. It's important to take into consideration how the different shapes impact how each prosthetic device operates and is perceived. An *anthropomorphic*, or human-like, prosthetic device will operate like its respective limb does.

Figure 3. A body–powered hook prosthetic [5].

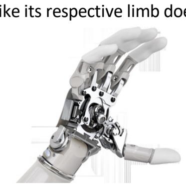


Figure 4. An anthropomorphic myoelectric prosthetic [6].

It can also increase the chance of the prosthetic device incorporating into how the amputee perceives or represents their body, a term known as *embodiment* [7]. Other factors, such as material, can contribute to a prosthetic device's form and function.

## Objectives

**Guiding Question:** Which thumb tip shape, when placed underneath a skin flap, is the most visually anthropomorphic?

**Goals of Design:**

- Must visually look as anthropomorphic as possible
- Must minimize the amount of extrusions

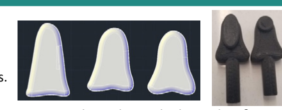
## Materials & Methods

1. Used AutoCAD to create 3D models of the thumb tip shapes, modeled after radiographic data of bone shape.
2. 3D printed each of the designs using Onyx filament.
3. Dissected a cadaver and placed the same radial forearm free flap around each of the designs.
4. Took photos of each thumb tip to decide which is the most anthropomorphic.

## Results & Discussion

Some of the previous 3D models are shown in Figure 5 in order to compare the past designs with the final 3D thumb tips used in the

Figure 5. Two previous designs.



cadaver dissection, shown in Figure 6 with their parameters listed out below the figure. The two shape types tested in the dissection were elliptical and conical, shown in Figure 6(a) and (b) and were modeled after radiographic data of the bone shape. All of the edges in the final models were rounded using the fillet tool in order to minimize the number of possible extrusions the device could cause.

### 3D Models


Figure 6: (a) Conical thumb tips, from smallest to biggest from left to right.
(b) Elliptical thumb tips of varying radii and widths.

The three conical shaped devices with varying radii are in Figure 6(a).

- The largest conical thumb tip has a bottom radius of 0.6 cm and a height of 2.2 cm
- The smallest one has a radius of 0.3 cm and a height of 1.9 cm
- The one in between has a radius of 0.45 cm and a height of 1.9 cm

Five elliptical shaped thumb tips are in Figure 6(b) with three different variations in the length and height of their ellipses.

- An ellipse of 1.8 cm high by 1.5 cm tall, or 1.5 cm long by 1 cm tall, or 1.4 cm long by 1.6 cm tall

The five elliptical shaped thumb tips had two variations in the width.

- 0.6 cm wide or 0.8 cm wide

### Dissection

In preparation for the dissection, cylinders of about 3 cm long were added under each thumb tip in order to represent the rest of the thumb for the sake of having more realistic photos, as seen in Figure 6. Fake paper thumbnails were created each thumb tip, as shown in Figure 7 to aid in making more realistic thumb tips.

Figure 7. A fake thumbnail created out of paper.

## Conclusion

Each of the final 3D models in Figure 6, when compared with their previous designs such as the ones in Figure 5, look more visually anthropomorphic and minimize the number of possible extrusions in the thumb tip due to the rounded edges and the symmetry in the prosthetic device. As a result, these thumb tips look more anthropomorphic in the skin flap transfer process, providing a way for embodiment to happen more easily.

### Next Steps

- Creating a survey
  - Gathering feedback and gauging the perception about each thumb tip in the dissection photos.

## References

[1] Yoon AP, Mahajani T, Hutton DW, Chung KC, for the Finger Replantation and Amputation Challenges in Assessing Impairment, Satisfaction, and Effectiveness (FRANCHISE) Group. Cost-effectiveness of Finger Replantation Compared With Revision Amputation. *JAMA Netw Open.* 2019;2(12):e1916509. doi:10.1001/jamanetworkopen.2019.16509

[2] Bregoli, Chiara, et al. "Osseointegrated Metallic Implants for Finger Amputees: A Review of the Literature." *Orthopaedic Surgery*, vol. 14, no. 6, 7 May 2022, pp. 1019–1033, https://doi.org/10.1111/os.13296.

[3] Karamanos, E., Julian, BQ., Cromack, D.T. (2021). The Radial Forearm Flap. In: Comprehensive Atlas of Upper and Lower Extremity Reconstruction. Springer, Cham. https://doi.org/10.1007/978-3-030-74232-4_13

[4] Lai, G. V. (2022). *Design and Validation of a Skin-Covered, Linkage-Driven Implantable Prosthesis* [Master's thesis, University of California, Los Angeles].

[5] "Split Hooks." *Steeper Group*, www.steepergroup.com/prosthetics/upper-limb-prosthetics/hands/split-hooks/.

[6] León, Beatriz, et al. "Robot grasping simulation." *Cognitive Systems Monographs*, 2013, pp. 33–65, https://doi.org/10.1007/978-3-319-01833-1_3.

[7] Zbinden, J., Lendaro, E. & Ortiz-Catalan, M. Prosthetic embodiment: systematic review on definitions, measures, and experimental paradigms. *J NeuroEngineering Rehabil* **19**, 37 (2022). https://doi.org/10.1186/s12984-022-01006-6

## Acknowledgements

I would like to thank the Samueli Research Scholars program for funding my research through the Summer Undergraduate Research Program (SURP) at UCLA. I would like to thank Professor Tyler Clites for allowing me to use his lab to research. I would also like to thank Sachi for supporting me and mentoring me throughout the summer as well as Dr. Lauren Wessel for performing the cadaver dissection. I would finally like to thank William Herrera and Samanta Negrete Munoz for helping to host SURP this summer.

**FACULTY ADVISOR**

**Timothy Fisher**

**DAILY LAB SUPERVISOR**

**Bryce Boyer**

**DEPARTMENT**

**Mechanical and Aerospace Engineering**


## ISABEL MARAVILLA

Mechanical and Aerospace
Engineering
2nd Year, UCLA


# Experimental Measurement of Work Function via Thermionic Emission of Molten Metals

## ABSTRACT


Ongoing research is considering the cooling of hypersonic vehicles via thermionic emission. Thermionic emission occurs when a material contains enough thermal energy that electrons escape from the surface. As high energy electrons escape from the surface, they are replaced with electrons of lower energy causing a net cooling effect. The work function must be known to accurately predict the cooling magnitude. While the work function for many materials is documented, emerging aerospace applications use newer materials or processes lacking a documented work function. In this project, we are creating an in-house system that can accurately measure the work function of materials. Our approach utilizes a high vacuum chamber with an electron energy analyzer, infrared camera, and thermocouples to measure the electron energy distribution from samples at elevated temperatures. The work function can be determined using the shape of the spectral emission intensity curve. Preliminary data has been used to verify proper operation of the equipment, and test procedures will be ready for full operation upon calibration with (100) tungsten in the coming weeks.

The significance of this research relates to improving heat management in hypersonic flight. Our work is a precursor to employing the High Energy Flux Test Facility (HEFTY) to measure cooling in a high heat flux environment. By determining the work function, we will be able to predict net cooling and validate modeling efforts with experimental data. This work lays the foundation for innovations in aerospace technology, with the potential to ultimately drive advancements in hypersonic flight capabilities.

SUMMER UNDERGRADUATE RESEARCH PROGRAM

# Experimental Measurement of Work Function via Thermionic Emission

Isabel Maravilla, Bryce Boyer, Prof. Timothy S. Fisher
Department of Mechanical and Aerospace Engineering, University of California, Los Angeles

## Objective

We are creating an in-house system that can accurately measure the work function of materials using thermionic emission spectrum at elevated temperatures.

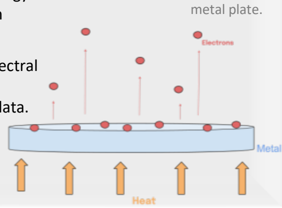
## Introduction

Ongoing research is considering cooling of hypersonic vehicles via thermionic emission. Thermionic emission occurs when electrons in surface layers of a material receive sufficient thermal energy to overcome the potential energy barrier of the material and escape. As high energy electrons escape from the surface, they are replaced with lower energy electrons causing a net cooling effect. Work function must be known to accurately predict cooling magnitude. While work function for many materials is documented, emerging aerospace applications use newer materials or processes lacking documented work function.

### Work Function

- Work function is minimum energy needed to remove an electron from the surface of the metal.
- Determined using shape of spectral emission intensity curve from energy emission distribution data.
- Highly dependent on surface properties and orientation of material.

Fig 1. Thermionic emission occurring in a metal plate.



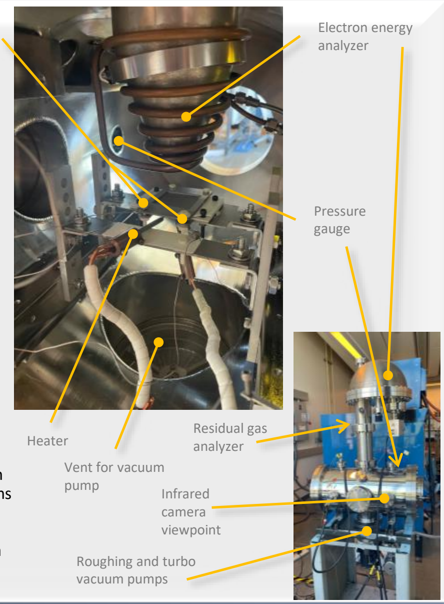

## Materials & Method

Calibration

- Work function of electron energy analyzer changes over time and must be accounted for to collect accurate emission distribution.
- (100) tungsten will be used to calibrate periodically.

Data Collection

- Data is collected via electron energy analyzer which measures distribution of emission energies.
- The analyzer works by creating an electric field between two plates with opposite charges. This allows electrons of certain energy to follow the path between the plates. Electrons that make it to the end are measured by a detector.




## Data Analysis

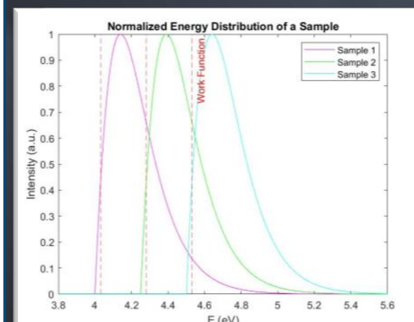
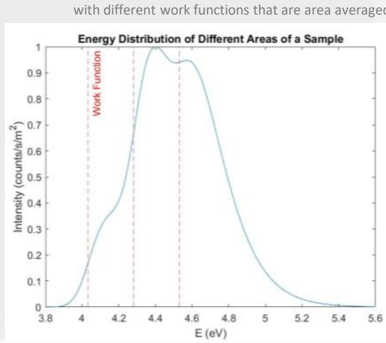



Fig. 2 A normalized theoretical energy distribution diagram for three different work functions marked by vertical dashed red lines.

- Work function can be found using an emission intensity curve using the following equation:
  $\phi = E_{Imax} - k_B T$
- Width of emission distribution is relative to temperature and can be calculated using $2.45k_B T$

- Work function is calculated using Python code that analyzes emission distribution from the electron energy analyzer.
- Fig 2 and 3 are theoretical models generated to compare our experimental results with. Fig 2 is a normalized plot for the same work functions in fig 3.

Fig. 3 Theoretical energy distribution diagram of one sample with different work functions that are area averaged.




## Future Work

Fig. 4 Partially assembled vacuum chamber with our High Energy Flux Test Facility (HEFTY) attached.

Our eventual goal is to measure cooling due to thermionic emission from the leading edge of materials used in aerospace vehicles. To do so, we have a hypersonic vacuum chamber. To simulate the high thermal effects of hypersonic environment, we use an arc jet system that provides a high heat flux.

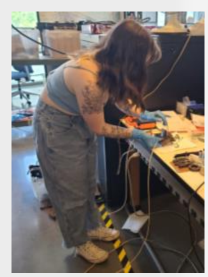


Fig. 5 Here I am in the lab working on a sample stage for HEFTY!

## Conclusion

Our project serves as a way for us to measure the work functions of materials in lab. The significance of this research relates to improving heat management in hypersonic flight. Our work is a precursor to employing the High Energy Flux Test Facility (HEFTY) to measure cooling in a high heat flux environment. By determining the work function, we will be able to predict net cooling and validate modeling efforts with experimental data. This work lays the foundation for innovations in aerospace technology, with the potential to ultimately drive advancements in hypersonic flight capabilities.

## References

[1] Westover, T. L. (2008). *Energy Transport and conversion in electron emission processes*.

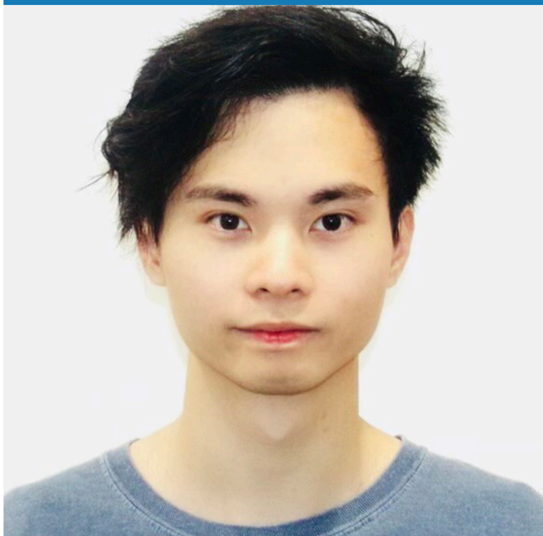
[2] Robinson, V. S. (2004). *An Experimental Study of Thermionic Emission Energy Distributions from Carbon-Based Material*.

[3] Fisher, T. S. (2014). *Thermal energy at the nanoscale*. World Scientific.

[4] Kwee, C. (2020, November 25). *Thermionic emission*. SPM Physics. https://spmphysics.blog.onlinetuition.com.my/electronic/thermionic-emission/

## Acknowledgements

I would like to thank my daily lab supervisor Bryce Boyer and P.I. Dr. Timothy S. Fisher for their constant mentorship and guidance. I would also like to thank the SURP program, specifically Will Herrera, Sam, Amir, and Zara, for all their help and support this summer. Special thanks to the Samueli Research Scholars for funding over the summer!




## JACOB SAYONO

Mechanical Engineering
4th Year, UCLA

**FACULTY ADVISOR**

**Yang Zhang**

**DAILY LAB SUPERVISOR**

**Xiaoying Yang**

**DEPARTMENT**

**Electrical and Computer Engineering**


# Interaction-Powered Light-Transfer Mechanisms for Ubiquitous Interactivity


## ABSTRACT

Retroreflectors, known for their widespread use in road signs and safety gear, serve a critical function in enhancing visibility. In this context, we explore their application in visible light communication (VLC), aiming to leverage their reflective properties for information transfer. Expanding on this concept, we devise mechanisms that encode distinct signal patterns generated by human interactions, thereby enabling self-sustaining smart sensing capabilities and creating dynamic interfaces for physical environments. By turning away from the reliance on conventional systems such as electronics or external power sources, our approach embeds relevant data seamlessly into the environment.

To evaluate this, we developed a mobile app that captures light intensities over time at a specified location within the camera view and investigated diverse fabrication methods related to retroreflective materials. At a small scale, 3D-printed remote controller mechanisms such as buttons, rotating knobs, and sliding switches validated our concept. At a medium scale, vinyl-cut retroreflective barcodes encoded information for indoor surfaces. At a large scale, CNC-machined arrays of triangular foam prisms used for city signage dynamically altered conveyed information based on viewing perspectives. Results demonstrate that our smart retroreflector design can enrich people's interaction with their surroundings, promoting efficient self-sustaining information distribution.

SUMMER UNDERGRADUATE RESEARCH PROGRAM

# Interaction-Powered Light Transfer Mechanisms for Ubiquitous Interactivity

**Jacob Sayono, Dr. Yang Zhang**

Department of Electrical and Computer Engineering, University of California – Los Angeles

## Introduction

Visible Light Communication (VLC) uses light within the visible spectrum to transmit information.

Retroreflectors reflect light back to their source regardless of the incoming angle.

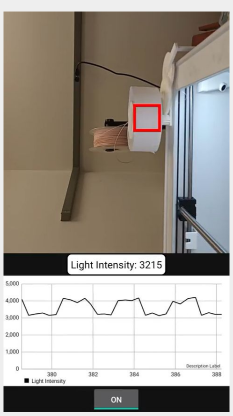
Interaction-Powered Systems derive functionality from the direct engagement of users.

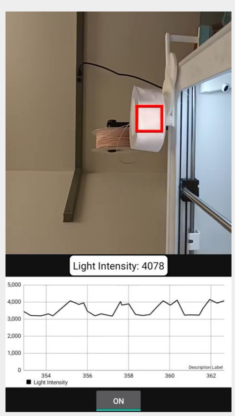
## Objective

We devise retroreflective mechanisms that encode distinct signal patterns generated by human interactions. This enables self-sustaining smart sensing capabilities and creates dynamic interfaces for physical environments through VLC.

## Implementation

A mobile app (developed using Java in Android Studio) captures light intensities in real-time at a specified location, indicated by a red bounding box. The app allows the user to turn on/off the phone flashlight and plots the intensity signals beneath the camera preview.



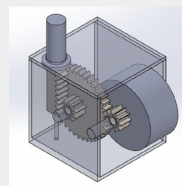



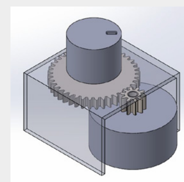

A 3D-printed mechanism, incorporating both retroreflective and non-retroreflective surfaces, can be attached to existing objects (e.g., a door). As users engage with the object, the mechanism exhibits the ability to reflect varying intensities of light over time.

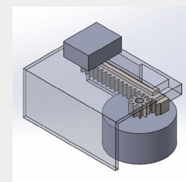
We explored a diverse set of applications, spanning from small to medium to city-scale, in order to showcase the versatility of our retroreflector-based smart sensing approach.

## Results

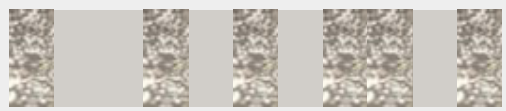
At a small scale, 3D-printed remote controller mechanisms such as buttons, rotating knobs, and sliding switches validated our concept by producing accurate signals.



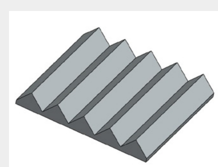


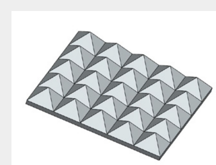


At a medium scale, vinyl-cut retroreflective barcodes encoded information onto surface appliances. Here, the retroreflective surface represents a bit of 1 while the non-retroreflective surface represents a bit of 0.



At a large scale, CNC-machined arrays of triangular prisms dynamically altered conveyed information based on viewing perspectives. This allows the sensor to locate its position relative to the retroreflectors.





## Conclusion

Our research innovatively combines retroreflectors and interaction-as-power to establish smart sensing. Our versatile approach across scales promises diverse applications, contributing to the evolving field of VLC and self-sustaining technologies at the intersection of digital and physical worlds.

## Acknowledgements

I extend my gratitude to my advisor Dr. Yang Zhang, the Human-Centered Computing & Intelligent Sensing Lab (HiLab), the Summer Undergraduate Research Program (SURP), and the National Science Foundation (NSF) for their roles in making this project possible.


**JEFFREY WENG**

Electrical and Computer Engineering
1st Year, UCLA


**FACULTY ADVISOR**

**Christina Fragouli**

**DAILY LAB SUPERVISOR**

**Mine Dogan**

**DEPARTMENT**

**Electrical and Computer Engineering**

# Implementation of Low Complexity Multi-Level Encoding Schemes for mmWave

## ABSTRACT


Within both the military and civilian context, millimeter-wave (mmWave) networks expand the available spectrum of transmission and provide high speed communication. However, a well-documented caveat to these networks is that they are highly sensitive to blockage. Thus, it is important to design resilient transmission mechanisms for the operation of these networks. In our work, we leverage rich multipath environments and accurate models that estimate link blockage probabilities to design low-complexity proactive transmission mechanisms for mmWave networks. In particular, in the Algorithmic Research in Network Information Flow lab, we have proposed low-complexity multi-level encoding schemes that achieve high information rate with low outage probability while proactively minimizing latency of the delay-sensitive communications.

My project for this summer has been to implement the proposed coding schemes over arbitrary mmWave networks, and to improve the proposed designs to balance the information rate with a graceful performance degradation. I implemented alternative coding schemes, and I evaluated the performance of the proposed coding scheme against alternative schemes. This offers critical insights on the viability and quantitative incentives of the proposed schemes in mmWave applications.

UCLA Samueli School of Engineering
SUMMER UNDERGRADUATE RESEARCH PROGRAM

# Multilevel Codes for mmWave Networks

Jeffrey Weng, Mine Dogan, Martina Cardone, Christina Fragouli
Algorithmic Research in Network Information Flow (ARNI) Lab
*Department of Electrical and Computer Engineering, University of California, Los Angeles*

## Introduction/Motivations

- Millimeter-wave (mmWave) networks offer high speed communication, and are used in a multitude of applications
- **Opportunities:**

V2X communications

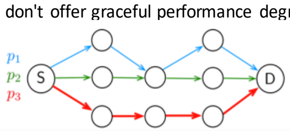
Terragraph by Meta

Unmanned aerial vehicles

- **Challenge**: mmWave links are highly sensitive to blockage
  - require resilient transmission mechanisms
- Proactive mechanisms build resilience ahead of time
  - offer low-delay communications

## Background: Coding Schemes

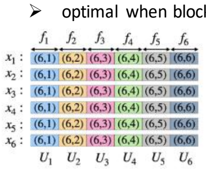
- Erasure correcting code: adds redundancy to offer reliability
  - don't offer graceful performance degradation

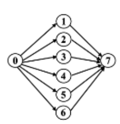



If we use (3,1) erasure correcting code and no link is blocked, we still experience rate 1/3!

## Equal Blockage Probabilities

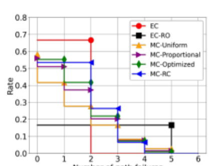
- Multilevel code combines multiple erasure codes for encoding source sequence
  - propose optimization problem to solve for packet fractions
  - optimal when blockage probabilities are uniform

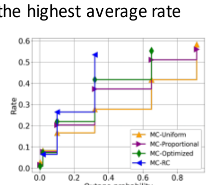


: An example network with 6 relays.

- Note: Complexity increases as the number of path blockages increases
  - Select $K$ codes with the highest average rate



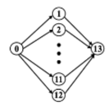




## References

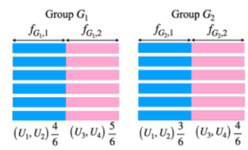
Dogan, M., Cardone, M. & Fragouli, C. Low-Delay Proactive Mechanisms for Resilient Communication

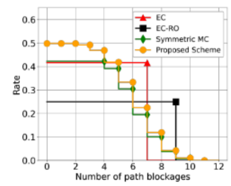
## Unequal Blockage Probabilities

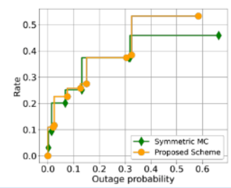
- Usually, blockage probabilities are not uniform
- Therefore, we propose an asymmetric multilevel code: divides paths into separate groups based on similarity in blockage probability
  - Assumes probabilities are not uniformly distributed
  - Clusters achieved rates depending on which erasure codes work successfully
  - Lowers complexity by using less erasure codes



An example network with 12 relays.

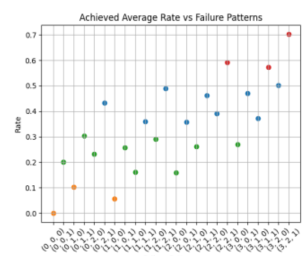



## Methods

- Python was used for entire implementation
- Libraries: CVXPY for solving optimization problems, Numpy for data processing and random sampling, scipy.stats for distribution sampling, sklearn for Kmeans clustering

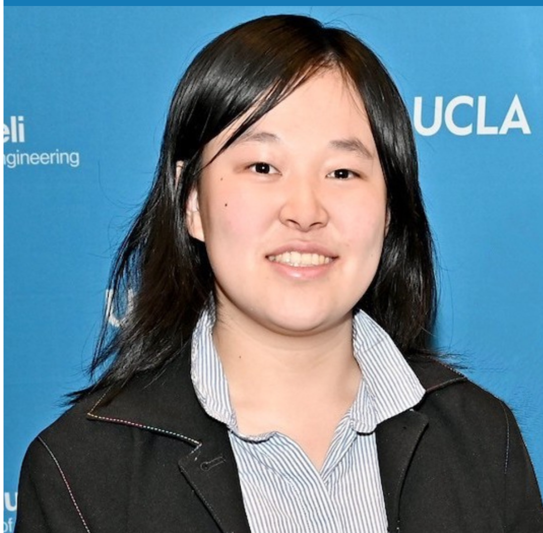



## Discussion

- Simulation program accepts arbitrary mmWave networks as parameters and compares rate-outage probability tradeoff
- Proposed asymmetric scheme outperforms alternative methods
- Multilevel code implementations are worth further exploration

## Acknowledgements

I would like to thank Professor Christina Fragouli and my daily lab supervisor PhD Mine Dogan for their continuous guidance during this program. I would also like to thank the UCLA Summer Undergraduate Research Program and the UCLA Department of Electrical and Computer Engineering for giving me this opportunity. This work was supported in part by the NSF grants 2229560, 2146838, and 2045237.




## FACULTY ADVISOR

**Amartya Banerjee**

## DAILY LAB SUPERVISOR

**Amartya Banerjee**

## DEPARTMENT

**Materials Science and Engineering**


# JENNIE REN

Materials Engineering
1st Year, UCLA


# First Principles Simulations of Excited States of Chiral Nanomaterials


## ABSTRACT

Chiral nanomaterials have been studied for their unique transport properties and potential to revolutionize electronics and quantum computing. A particularly important phenomenon is the chiral induced spin-selectivity effect (CISS), in which electrons are selectively transported through a material depending on their spin and the material's chirality. However, CISS and related phenomena are difficult to model using current first principles or quantum mechanics-based simulations as they involve complex systems of evolving electrons, spin, and geometry. Therefore, to design an accurate computational framework, we have implemented a time-dependent variant of Kohn-Sham Density Functional Theory (DFT), suited to chiral structures. Kohn-Sham DFT is the most widely used ab initio electronic structure calculation method.

We have carried out our implementation in multiple steps. First, we built upon the pre-existing finite difference Kohn-Sham DFT code M-SPARC, which is only capable of modeling ground state behavior for simple molecular geometries and obtained a time dependent version of the code. We accomplished this by utilizing numerical integration methods such as Runge Kutta to propagate the electron density over time. Then, we adopted the resulting methods to Helical DFT, a finite difference DFT code for helical (chiral) nanostructures. Helical DFT uses a helical coordinate system (instead of the rectangular one used in M-SPARC), allowing us to account for the twisted structure of chiral nanomaterials. We anticipate that our model will allow for better understanding of the electronic properties of these materials and how to harness their properties for constructing quantum devices.

UCLA Samueli School of Engineering

SUMMER UNDERGRADUATE RESEARCH PROGRAM

# First Principles Simulations of Excited States of Chiral Nanomaterials

**Jennie Ren, Prof. Amartya Banerjee**
Ab Initio Simulations Lab
Department of Materials Science and Engineering - University of California, Los Angeles

## Introduction

First principles simulations model materials by solving the **Schrödinger Equation**, which becomes impossible for large, complex systems. Thus, we use approximation methods such as Kohn-Sham **Density Functional Theory (DFT)**, but these typically only model **ground state** systems.

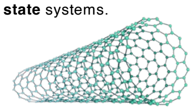


**Figure 1.** Illustration of a chiral carbon nanotube.

**Equation 1.** The Time-Dependent Schrödinger Equation:

$$i\hbar\frac{d}{dt}|\psi(t)\rangle = \hat{H}|\psi(t)\rangle$$

**Chiral nanomaterials**, due to their twisted structures, may revolutionize quantum computing and exhibit novel transport properties such as the **Chiral Induced-Spin Selectivity Effect**. We want to create an accurate first principles framework to model these unique materials and their **excited states**.

## Objective

Implement first principles simulation models in MATLAB to study chiral nanomaterials and their excited state behavior

## Technical Terms

**Chiral:** a chemical species with a mirror image that is non-superimposable upon itself

**Electron Density:** probability of finding electrons in a certain location around an atom or molecule

**External "Kick" or Pulse:** a temporary disturbance from an external source that changes a system's properties

**Ground State:** electron configuration with the lowest energy, any other arrangement is an **excited state**

## Materials

MATLAB
HP Laptop
M-SPARC: MATLAB Simulation Package for Ab Initio Real-Space Calculations

## Methods

### 1 Understand M-SPARC and its limitations

Understand the design of M-SPARC, a finite difference, ground-state DFT code and ensure our setup is correct by running tests on a sample **hydrogen molecule**, choosing an appropriate mesh size and creating electron density contour plots

### 2 Implement a time-dependent DFT framework

Build upon M-SPARC and use numerical integration methods such as **Euler's method** and **Runge Kutta** to propagate the system over time

**Equation 2.** Euler method equation for propagating the wavefunction over time in Hartree atomic units:

$$|\psi(t+\Delta t)\rangle = |\psi(t)\rangle - i\,\Delta t\,\hat{H}[n]|\psi(t)\rangle$$

**Equation 3.** Second order Runge Kutta equations for propagating the wavefunction over time in Hartree atomic units:

$$|k_1\rangle = -i\,\Delta t\,\hat{H}[n_{\psi(t)}]|\psi(t)\rangle$$

$$|k_2\rangle = -i\,\Delta t\,\hat{H}[n_{\psi(t)+0.5k_1}]|\psi(t)+0.5k_1\rangle$$

$$|\psi(t+\Delta t)\rangle = |\psi(t)+k_2\rangle$$

### 3 Adopt resulting methods to Helical DFT

Apply time-dependent DFT code to our pre-existing Helical DFT simulation framework, which allows us to account for twisted geometries

## Results

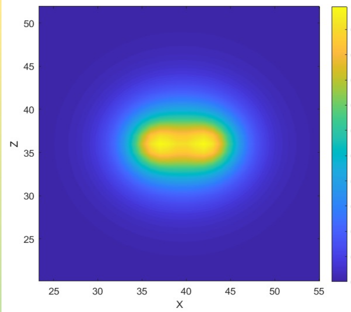



**Figure 2.** Electron density plot of a ground state hydrogen molecule. The two hydrogen atoms, around which most of the density is located, are clearly visible.

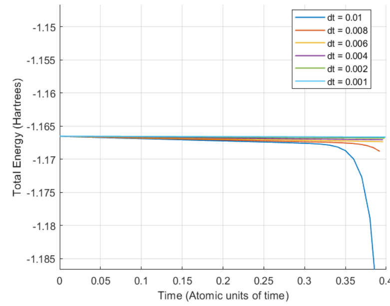



**Figure 3.** Plot of the total energy versus time for different time intervals (dt) when no external potential is applied for the Euler method code.

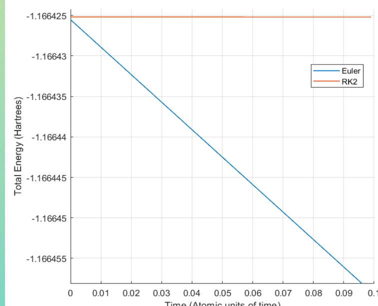



**Figure 4.** Plot of the total energy versus time as a measure of stability for both the Euler method and second order Runge Kutta (RK2) code in the absence of an external kick.

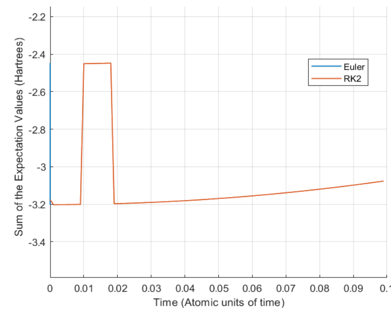



**Figure 5.** Plot of the sum of the expectation values versus time for Euler method and second order Runge Kutta schemes when an external kick of 0.001 Hartrees is applied from 0.01 to 0.02 atomic units of time.

## Discussion

In the absence of an external kick, the number of **fluctuations** in the total energy decreases with smaller values of dt (Figure 3). This matches the results of past studies and indicates a functioning and stable time-dependent DFT code. Second order Runge Kutta was also found to be more accurate than the Euler scheme, as the Euler scheme **drifts** more rapidly (Figure 4).

**Equation 4.** Sum of the expectation values for each wavefunction *i*, a measure of the average energy:

$$\sum_i \langle\psi_i(t)|\hat{H}[n]|\psi_i(t)\rangle$$

For the Euler method code, the **sum of the expectation values** versus time curve immediately diverges and displays unstable behavior when an external pulse is applied. For the second order Runge Kutta code, the curve exhibits a noticeable jump during the duration of the pulse before stabilizing at the original energy, which agrees with theory. However, it begins to drift as the simulation continues forward in time (Figure 5). This indicates that a more accurate numerical integration scheme may be needed.

### Next Steps

Create a more accurate time-dependent DFT framework using **fourth order Runge Kutta**
Find a more appropriate kick to accurately simulate an external energy source, such as a laser
Finish applying time-dependent DFT to Helical DFT code to study chiral nanomaterials

## References

Schleife, André, et al. "Plane-wave pseudopotential implementation of explicit integrators for time-dependent Kohn-Sham equations in large-scale simulations." *The Journal of Chemical Physics*, vol. 137, no. 22, 22 Oct. 2012. *AIP Publishing*, doi.org/10.1063/1.4758792.
Yu, Hsuan Ming, and Amartya S. Banerjee. "Density functional theory method for twisted geometries with application to torsional deformations in group-IV nanotubes." *Journal of Computational Physics*, vol. 456, 1 May 2022. *ScienceDirect*, doi.org/10.1016/j.jcp.2022.111023.
Zhang, Boqin, et al. "Version 2.0.0 - M-SPARC: Matlab-Simulation Package for Ab-initio Real-space Calculations." *SoftwareX*, vol. 21, no. 101295, 4 Jan. 2023. doi.org/10.1016/j.softx.2022.101295.

## Acknowledgements

I would like to thank Professor Amartya Banerjee and the Ab Initio Simulations Lab for their guidance and support. I would also like to thank the Samueli Research Scholars program for funding this project as well as William Herrera and the Summer Undergraduate Research Program staff for organizing this program.


## JESS XU

Computer Science
3rd Year, UCLA

**FACULTY ADVISOR**
**Yang Zhang**

**DAILY LAB SUPERVISOR**
**Siyou Pei**

**DEPARTMENT**
**Electrical and Computer Engineering**


# Mechanical Energy Harvesting Software Toolkit

## ABSTRACT


Mechanical energy harvesting is a promising way to power devices with low power requirements, such as sensors and microcontrollers. For example, previous work, such as MiniKers, has found success in harvesting mechanical energy from manual uses of household objects to provide energy for automatic actuations. However, selecting and characterizing motors for energy-harvesting circuits often involves a long process of trial-and-error, because the force and speed used to actuate a motor can vary greatly between different people and situations, and it can be difficult to estimate the amount of energy that can be harvested from motor datasheets alone.

The objective of this project is to create a novice-friendly desktop application that allows users to build energy-harvesting circuits and select parameters for manual actuation for simulation. Our tool builds on Fritzing, an open-source circuit schematic editor, which includes a breadboard view to allow users to visualize circuits similar to how they would build them in real life, for the circuit-building step. We extend Fritzing's simulation feature using ngspice to provide current and voltage waveforms and allow users to estimate the net amount of energy that a circuit can harvest based on motor model, gear ratio, RPM, and circuit components.

# Mechanical Energy Harvesting Software Toolkit

Jess Xu, Professor Yang Zhang
Human-Centered Computing & Intelligent Sensing Lab
Department of Electrical and Computer Engineering, University of California - Los Angeles
Collaboration with Ka Moamoa Lab (Georgia Institute of Technology)

UCLA Samueli School of Engineering

SUMMER UNDERGRADUATE RESEARCH PROGRAM

## Background and Motivation

Mechanical energy harvesting leverages the manual rotation of motors to store energy for later use. This energy can power devices such as microcontrollers, sensors, flashlights, and radios. Previous work such as MiniKers[1] has found success in harvesting mechanical energy from manual uses of objects to provide energy for automatic actuations.

In building energy harvesting circuits, selecting a suitable motor, gear ratio, and resistor can be confusing and time-consuming for novices, which motivates a software application that can facilitate the prototyping process.

## Objective

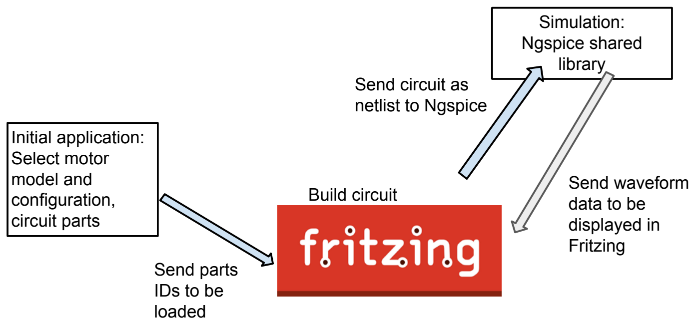
The objective is to create a novice-friendly desktop application that allows users to build and simulate their own energy harvesting circuits. Required features include allowing users to select parts for their circuit (and specify parameters such as motor model, RPM, and number of revolutions), a breadboard view that allows users to visualize circuits similar to how they would build them in real life, and a simulation feature that allows users to view resultant current and voltage waveforms and estimate the amount of energy harvested, as well as the amount of energy consumed by circuit components and the user's selected microcontroller (based on the microcontroller model and the program that it runs)

## Application Architecture

The circuit-building and simulation step is an extension of Fritzing, an open-source circuit schematic editor, which is written in C++ using the graphical user interface (GUI) library Qt. It uses Ngspice's shared library to perform simulations. We added and modified SPICE models of existing Fritzing parts, most notably a motor and bridge rectifier.

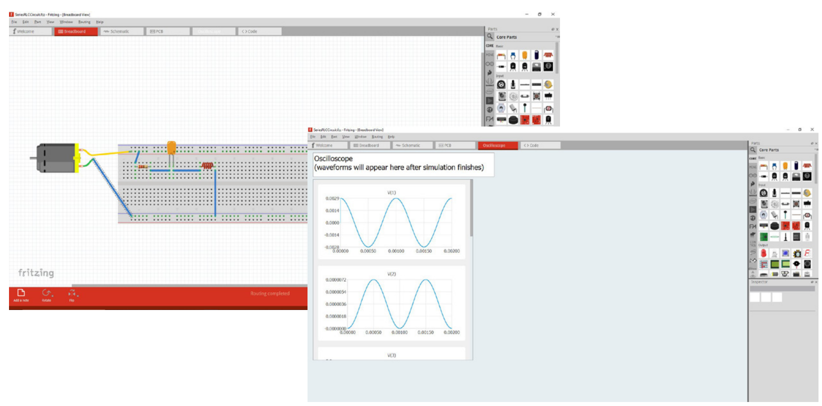
The initial application, where users select their circuit parts, is written in C++ using Qt as the GUI library. It accesses a database of circuit parts, including motors, that the user can choose from. When it starts the Fritzing process, it passes the parts IDs and motor configuration as arguments.

Ngspice stores current and voltage waveform data as *vectors*. When the simulation is complete, our application retrieves all of them and displays them in the oscilloscope tab.




## Interaction Flow

1. Application sends user-selected motor model and parameters to a modified version of Fritzing, over a TCP/IP socket.
2. User then builds their circuit inside Fritzing. When they are finished, they click a button in Fritzing.
3. A motor crank appears onscreen, and the user has the option to spin it as they would in real life. The user can choose to use an existing voltage waveform instead of cranking the motor.
4. Fritzing exports the user's sketch to a SPICE netlist. The simulation runs in a background thread, with the user-selected motor parameters being used to customize a PWL input to the simulated motor.
5. After simulation finishes, user can view waveforms of current and voltage in a "oscilloscope" tab in Fritzing.




## Results & Discussion

Currently, the app achieves the basic functions of allowing the user to select parts, build a circuit, and display the results of the circuit simulation. Future work includes making the application more flexible so that users can add motor models and configurations, as the current selection is limited to the existing ones in the database. We would also like to add a feature that allows the user to probe nodes on the breadboard, similar to an oscilloscope or multimeter, to allow them to measure particular voltages and currents (the app currently displays every branch current, node voltage, and differential voltage across predefined loads).

The application currently calculates the power harvested from the actuation of the motor. One goal is to calculate the power consumed by the user's selected microcontroller, taking into account the baseline power consumption of the particular microcontroller model and the code uploaded to it (e.g. more frequent Bluetooth advertisements result in higher power consumption)

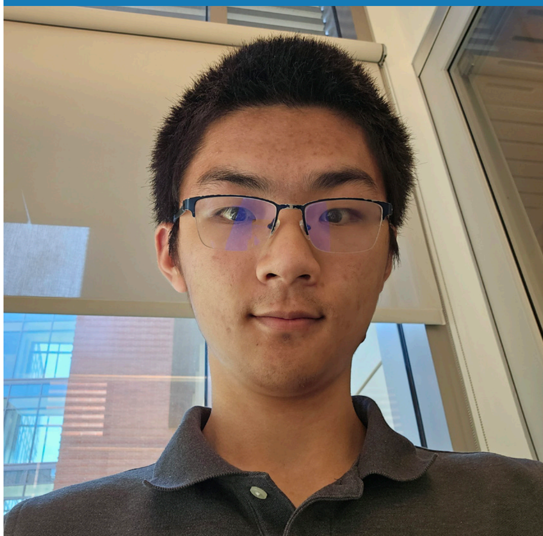
This application is part of a longer-term project, and we plan to conduct a user study on the prototype to determine how to make the tool more user-friendly.

## Acknowledgements

This research was funded by a grant from the National Science foundation. I would like to thank Dr. Yang Zhang, Dr. Josiah Hester, Xiaoying Yang, and Sydney Young for their guidance, knowledge and support throughout the project. I would also like to thank the Summer Undergraduate Research Program staff for facilitating this project.

## References

[1] Xiaoying Yang, Jacob Sayono, Jess Xu, Jiahao Nick Li, Josiah Hester, and Yang Zhang. 2022. MiniKers: Interaction-Powered Smart Environment Automation. Proc. ACM Interact. Mob. Wearable Ubiquitous Technol. 6, 3, Article 149 (September 2022), 22 pages. https://doi.org/10.1145/3550287



**FACULTY ADVISOR**

**Richard Wesel**

**DAILY LAB SUPERVISOR**

**Semira Galijasevic**

**DEPARTMENT**

**Electrical and Computer Engineering**


## JINGCHAO LUO

Electrical and Computer Engineering
1st Year, UCLA


# Maximizing Throughput over a Fading Optical Channel


## ABSTRACT

The Free Space Optical (FSO) Channel allows for communication across a much larger bandwidth compared to RF. However, atmospheric interference (fading) such as clouds or dust can negatively impact the reliability of the signal. Low Density Parity Check (LDPC) Codes can correct transmission errors, but their effectiveness is diminished if the fades are too strong. By lowering the LDPC Code Rate, fewer information bits are transmitted per second, but the signal becomes more resistant to fading.

This project investigates a method to maximize throughput over an FSO Channel by dynamically altering the LDPC Code Rate based on the current strength of fading. Thresholds for switching between 8/9, 8/10, … 8/16 code rates were precomputed by finding the average signal to noise ratio (SNR) that results in a Frame Error Rate of 10-2. A generated signal was simulated over a fading channel, with each transmitted block operating at the rate determined by the SNR thresholds. By always adapting the code rate to the highest threshold that matches the current fading strength, throughput increased by 33% compared to if only the single 8/9 rate was used. Further tests showed that adding an extra 0.25dB margin to the thresholds struck an even better balance between speed and reliability, resulting in a further 5.9% increase in signal throughput.

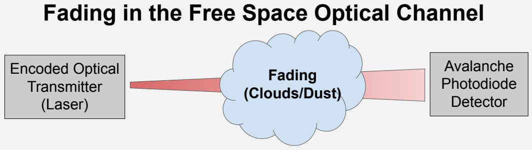
# Maximizing Throughput over a Fading Optical Channel

**Jingchao Luo, Semira Galijasevic, and Richard Wesel**
**Department of Electrical and Computer Engineering, UCLA**

UCLA Samueli School of Engineering
SUMMER UNDERGRADUATE RESEARCH PROGRAM

## Introduction

### Fading in the Free Space Optical Channel




- FSO Channel on-off keying modulates **2.5 Gbps** of LDPC bits [1]
- Compromised by light scattering interference (clouds, dust, rain)

- Use Gaussian distributions to model signal vs background noise [2]
- Bit 1 modulated to $\mu_s$ and passed through AWGN fading channel

$$\text{Detected Signal when bit 1 is sent: } y \sim \mathcal{N}(\mu_s, \sigma_s^2)$$
$$\text{Detected Signal when bit 0 is sent: } y \sim \mathcal{N}(\mu_b, \sigma_b^2)$$

### LDPC Codes

- Efficient encoding technique makes signal resilient to noise
- Utilizes sparse Tanner graph, parity and generator matrices [3]

$$\text{Code Rate } R_C = \frac{k \text{ message bits}}{n \text{ transmitted bits}}$$

- Each transmitted codeword has message bits + "redundancy" bits
- Trade-off between speed and fading resilience

### Bit Log Likelihood Ratio (LLR)

- Used for LDPC decoding
- Values of $\mu_s$, $\mu_b$, $\sigma_s$, $\sigma_b$ evolve each codeword based on fading model

$$\lambda_b = \frac{1}{2}\ln\frac{\sigma_b^2}{\sigma_s^2} + \frac{(y_j - \mu_b)^2}{2\sigma_b^2} - \frac{(y_j - \mu_s)^2}{2\sigma_s^2}$$

## Methods

1. Generated codewords are sent through an unfaded optical channel with a baseline power of -53.9 dB (power detected from background noise).

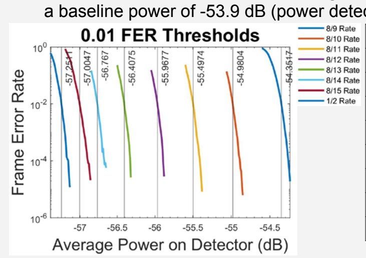


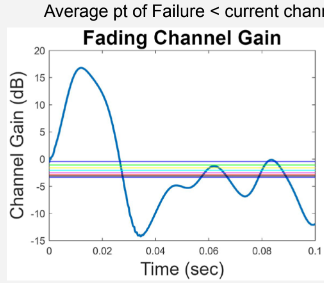

Figure 1. FER Waterfall Curves vs Power on Detector

| Code Rate | Avg pt of failure |
|---|---|
| 8/9 | -0.435 dB |
| 8/10 | -1.080 dB |
| 8/11 | -1.597 dB |
| 8/12 | -2.068 dB |
| 8/13 | -2.508 dB |
| 8/14 | -2.867 dB |
| 8/15 | -3.105 dB |
| 8/16 | -3.351 dB |

Table 1. Avg decoding failure pt power thresholds

$$\text{Average pt of Failure} = (\text{Power at 0.01 FER}) - (\text{Baseline Power})$$

2. Signal is passed through a deep fading optical channel, with code rate for each individual block chosen as the highest rate with its corresponding Average pt of Failure < current channel gain.



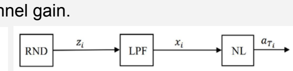




Figure 4. Fading channel realization is created by passing a Gaussian distribution of values through a low-pass filter and a non-linear transform [2].

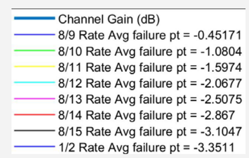



Figure 3. Fading Channel Realization. For any given rate, the chance of a successful transmission heavily depends on if the current fading gain falls above or below the Average failure pt Line.

## Objective

This project aims to achieve the highest possible throughput over a deep fading channel. By dynamically adapting the LDPC code rate to the current strength of fading being experienced, the transceiver system maintains a balance between signal speed and reliability.

## Materials

- The LDPC codes used in this project, including the involved puncturing pattern and generator matrix, were developed by Chen, et al. [3] at UCLA.
- The simulations were run in C++ using source code adapted from Nguyen, et al. [2], and the results plotted in Matlab.

## Results

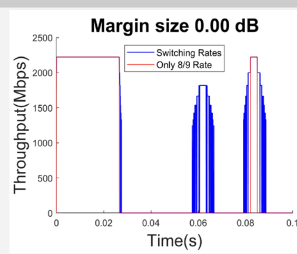



Figure 5. Realtime Throughput(Original Thresholds)

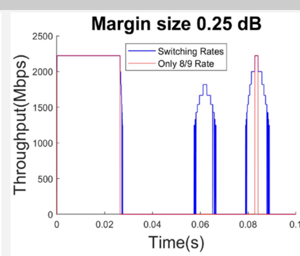



Figure 6. Realtime Throughput(Increased Thresholds)

| Signal Statistics | Margin Size = 0.00 dB | | Margin Size = 0.25 dB | |
|---|---|---|---|---|
| | Switching Rates | Only 8/9 Rate | Switching Rates | Only 8/9 Rate |
| **Throughput Over 0.1s** | 87.29 Mb | 65.48 Mb | 92.44 Mb | 61.52 Mb |
| **Decoding Failure Rate** | 7.71% | 0.00% | 0.17% | 0.00% |

Table 2. Margin size increases the power on detector threshold for selecting each code rate, lowering the Decoding Failure Rate - "variance" based transmission failures at the boundary of each code rate.

## Conclusions

- Dynamically adapting the LDPC Code Rate significantly increases overall throughput by **33%**
- Lower Code Rates are effective at maintaining a stable transmission link, at an albeit lower speed, when the Fading Channel Gain is between 0 and -3 dB – seen here from 0.05s to 0.09s.
- Including a 0.25 dB margin to the Code Rate thresholds drops the number of decoding failures in this 0.1s time span from 896 to 27, resulting in an overall throughput increase of **5.9%**

## Acknowledgements

This research is supported by the UCLA Communication Systems Laboratory, CACI and National Science Foundation Grant CCF-1955660.
I am extremely grateful for Prof. Richard Wesel and Semira Galijasevic for their continued guidance and support. I would also like to thank Jonathan Nguyen for providing documentation for the source code used in this project. And finally, I would like to acknowledge that this project used computational and storage resources associated with the Hoffman2 Shared Cluster provided by UCLA Office of Advanced Research Computing's Research Technology Group.

## References

[1] M. W. Wright, M. Srinivasan and K. Wilson, Improved Optical Communications Performance Using Adaptive Optics with an Avalanche Photodiode Detector, vol. 42, 2005.
[2] J. Nguyen, E. M. Liang, L. Wang, R. D. Wesel, T. Drullinger and T. Chauvin, "Comparison of Integrated and Independent RF/FSO Transceivers on a Fading Optical Channel," 2020 54th Asilomar Conference on Signals, Systems, and Computers, Pacific Grove, CA, USA, 2020, pp. 699-701, doi: 10.1109/IEEECONF51394.2020.9443554.
[3] T.-Y. Chen, K. Vakilinia, D. Divsalar and R. Wesel, "Protograph-Based Raptor-Like LDPC Codes," IEEE Transactions on Communications, vol. 63, no. 5, pp. 1522-1532, May 2015.

## FACULTY ADVISOR

**Nader Sehatbakhsh**

## DAILY LAB SUPERVISOR

**Pooya Aghanoury**

## DEPARTMENT

**Electrical and Computer Engineering**


# JOSEPH SEOK

Electrical and Computer Engineering
1st Year, UCLA


# CPU Side-Channel Fingerprinting


## ABSTRACT

This research delves into device fingerprinting using CPU side-channel emanations, capturing the innate information leakage of these channels. Through tools like DeMiCPU, which collects magnetic induction signals from CPUs, we have unlocked a method for deriving distinct device signatures, critical for enhanced security in both software and applications. Central to our approach is a deep learning-based classifier that capitalizes on side-channel data from microcontroller CPUs. Raw IQ data is gathered with an EM probe, and, after undergoing FFT for frequency domain analysis and normalization, is ready for model training. We implemented both a Convolutional Neural Network (CNN) for image-based spectrogram data and a Long Short-Term Memory (LSTM) model for sequence-based IQ data.

Our methodology involved recording EM signals from four Arduino devices at several different time instances, resulting in an extensive FFT dataset. Though the LSTM model initially performed well in identifying trained devices, it misidentified an untrained device as a trained device with high confidence, hinting at overfitting. However, by revising the normalization function and introducing an improved validation prediction rule, we achieved 100% accuracy in identifying both trained and untrained devices. In conclusion, this project explores an accurate method of utilizing the CPU's EM emanations for device fingerprinting. By merging FFT-based frequency domain analysis with refined deep learning techniques, we offer a method that suggests potential advancements in device identification, contributing to the field of cybersecurity.

# CPU Side-Channel Fingerprinting

Joseph Seok, Pooya Aghanoury, and Nader Sehatbakhsh
Department of Electrical and Computer Engineering, University of California – Los Angeles

UCLA Samueli School of Engineering
SUMMER UNDERGRADUATE RESEARCH PROGRAM

## Introduction

•Our research explores the use of CPU side-channel electromagnetic (EM) emanations as a means for device fingerprinting.
•This approach aims to enhance device authentication methods, highlighting the potential importance of electromagnetic signals in cybersecurity.

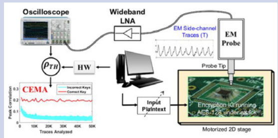



**Figure 1:** EM emanations can be traced and analyzed.

## Project Overview

•We design a Deep Learning-based classifier using side-channel information from microcontroller CPU's.
  •Side-channel Information : Raw IQ data collected from EM probe.
  •Data preprocessing : FFT to analyze IQ in frequency domain.
  •Deep learning modeling : =
    •Convolutional neural network (CNN) model.
    •Long short-term memory (LSTM) model.
  •Performance evaluation :
    •Trained devices are identified as trusted.
    •Untrained devices are identified as untrusted.

## Materials

•Our methodology recorded EM signals from four Arduino devices at various times, resulting in an extensive FFT dataset.

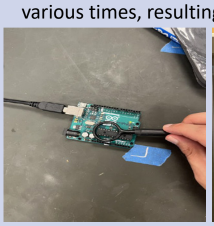


**Figure 2:** An idle Arduino being probed near the CPU.

**Figure 3:** Spectrum Analyzer screen displaying spikes in CPU emanations from clock signals.

## Procedures

•EM signals from Arduino devices 1-4 were measured five times at 600KHz, each generating 1249 x 1024-point FFT data.
  •First set of signals from devices 1-3 were used to train an LSTM deep learning model.
  •The next four sets of signals tested the model's classification abilities.
  •Signals from device 4 tested the model's ability to identify unclassified signals.

## Deep Learning Modeling

•Our LSTM-based model classifies EM signals from Arduino devices with the following structure:
  •Input: 1024-length EM signal sequences.
  •LSTM: 50 units; captures patterns, outputs final value.
  •Connected Layer: 3 nodes for each Arduino device.
  •Softmax: Calculates device probability.
  •Classification: Predicts device from softmax.
•Using 'adam' with validation, the model recognizes signals and tests adaptability.

```
%% Build the LSTM Model:
layers = [ ...
    sequenceInputLayer(1024)
    lstmLayer(50, 'OutputMode', 'last')
    fullyConnectedLayer(3) % 3 classes corresponding to 3 devices
    softmaxLayer
    classificationLayer];
```

```
%% Training Options:
options = trainingOptions('adam', ...
    'MaxEpochs',10, ...
    'MiniBatchSize',32, ...
    'Shuffle','every-epoch', ...
    'ValidationData',{X_test_cell, y_test}, ...
    'ValidationFrequency',30, ...
    'Verbose',false, ...
    'Plots','training-progress');
```

**Figure 4:** LSTM modeling in MATLAB

## Results

•Perform a 1024-point FFT on I/Q samples of four Arduino devices.
•Train Arduino devices 1 through 3 using an LSTM model.
•Test Arduino 4 with the trained LSTM model.
•**Discussion 1**: Model shows 25% accuracy and therefore fails to identify devices.
  •Proposal: LSTM model's normalization function is changed from dB scale to Min-Max scaling.

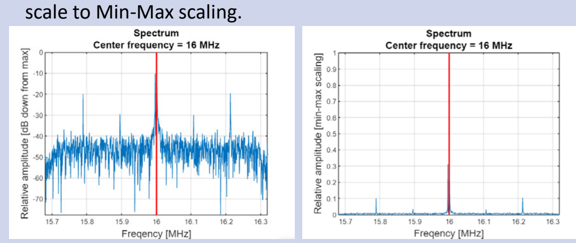



**Figure 5:** dB Scaling vs. Max-Min Scaling.

•**Discussion2:** Model shows 100% accuracy but fails to identify unclassified devices.
  •Proposal: 20% of the sample is replaced with an incorrect signal to address overfitting. Validation rule updated to:
  *validThreshold = (YPred > MinConfidenceLevel) &*
  *(YPred < MaxConfidenceLevel);*

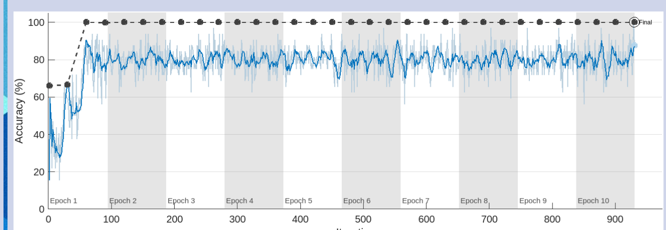



**Figure 6:** LSTM Training Considering Unclassified Devices

•**Model achieves 100% accuracy and can identify unclassified devices.**

```
Train accuracy: 100%
Test with a new electromagnetic signal obtained from trained device:
The new sequence is similar to device 1
The new sequence is similar to device 1
The new sequence is similar to device 1
The new sequence is similar to device 1
The new sequence is similar to device 2
The new sequence is similar to device 2
The new sequence is similar to device 2
The new sequence is similar to device 2
The new sequence is similar to device 3
The new sequence is similar to device 3
The new sequence is similar to device 3
The new sequence is similar to device 3
Test with a new electromagnetic signal obtained from non-trained device:
The new sequence is similar to device 2
The new sequence is similar to device 2
The new sequence is similar to device 2
The new sequence is similar to device 2
Test with a random sequence:
The new sequence is anomalous.
The new sequence is anomalous.
The new sequence is anomalous.
The new sequence is anomalous.
```

```
Train accuracy: 100%
Test with a new electromagnetic signal obtained from trained device:
The new sequence is similar to device 1
The new sequence is similar to device 1
The new sequence is similar to device 1
The new sequence is similar to device 1
The new sequence is similar to device 2
The new sequence is similar to device 2
The new sequence is similar to device 2
The new sequence is similar to device 2
The new sequence is similar to device 3
The new sequence is similar to device 3
The new sequence is similar to device 3
The new sequence is similar to device 3
Test with a new electromagnetic signal obtained from non-trained device:
The new sequence is anomalous.
The new sequence is anomalous.
The new sequence is anomalous.
The new sequence is anomalous.
Test with a random sequence:
The new sequence is anomalous.
The new sequence is anomalous.
The new sequence is anomalous.
The new sequence is anomalous.
```

**Figure 7:** LSTM Modeling Testing Result Comparison

## Conclusion

•In our study, we explored using EM outputs from the CPU's clock for RF fingerprinting. Using FFT, we accurately identified devices. Future research might apply this for authentication, location, and secure communication.

## Acknowledgements and References

•This work has been supported, in part, by NSF grant CNS-2211301. The views and findings in this poster are those of the authors and do not necessarily reflect the views of NSF. Special thanks to Prof. Sehatbakhsh for his knowledge and support as well as the opportunity to work in the Secure Systems and Architectures Lab.
•Hertz, Jake. "Em Side-Channel Attacks on Cryptography." All About Circuits, 26 July 2023, www.allaboutcircuits.com/technical-articles/em-side-channel-attacks-on-cryptography/.
•Ji, Xiaoyu et al. "Device Fingerprinting with Magnetic Induction Signals Radiated by CPU Modules." ACM Transactions on Sensor Networks (TOSN) 18 (2021): 1 - 28.


FACULTY ADVISOR

**Sudhakar Pamarti**

DAILY LAB SUPERVISOR

**Haris Suhail and Vinod Jacob**

DEPARTMENT

**Electrical and Computer Engineering**

## JOSHUA BURPEE

Electrical and Computer
Engineering
1st Year, UCLA


# Design and Test a PCB for a Magnetic Random Access Memory Compute-in-Memory Chip


## ABSTRACT

Advances in computing technology have brought us to thousands of calculations performed in milliseconds, however moving data from memory to processor takes substantial time and energy. This begs the need to rethink data storage in high-performance systems to eliminate this bottleneck, particularly in machine learning workloads where millions of operands and parameters are used. With traditional memory candidates like SRAM, eDRAM, and Flash reaching physical limits as technology evolves, Magnetic Random Access Memory (MRAM) is a promising solution. By storing information using magnetic states, MRAM is non-volatile and illustrates the potential of compact, low power data storage scalable with technology.

Voltage Control Magnetic Tunnel Junctions (VC-MTJs) are one demonstration of MRAM, having less power consumption and higher resistance, allowing for efficient read and write of the device. Our IC chip combines numerous VC-MTJ devices, utilizing CMOS fabrication, into four separate memory arrays. The IC is designed to perform efficient multiply-and-accumulate (MAC) operations which comprise upwards of 80% of all neural network computations. By computing locally with information stored in the memory, rather than transporting that data to a separate processor, this IC creates compute-in-memory. My project is to design a PCB that will interface with this chip to test the limits of reading, writing, and computing. Starting from scratch, the PCB has independent power supplies to operate each compute-in-memory array separately, a simplistic way to read the analog output signals, and constant current and voltage references. Once manufactured, the PCB will be tested using JTAG serial interfacing and python-controlled scripts to confirm this promising potential.

# Design and Test a PCB for a Magnetic Random Access Memory Compute In Memory Chip

UCLA Samueli School of Engineering
SUMMER UNDERGRADUATE RESEARCH PROGRAM

Joshua Burpee, Vinod Kurian Jacob, Haris Suhail, Prof. Sudhakar Pamarti
Department of Electrical and Computer Engineering, University of California, Los Angeles

## Introduction

- Traditional computing architecture can not keep up with increasing demands in machine learning
- Memory access dominates compute latency creating the immediate demand for a redesign of computer architecture
- The solution is a chip that is capable of storing and performing computations on one device, hence compute-in-memory
- Magnetic memory devices have shown substantial potential yet current renditions are power inefficient and vary between devices

## Key Terms

- MRAM – Magnetic Random Access Memory
  - Stores data as two different magnetic states
  - Non volatile memory that holds information even after losing power
- VC-MTJ – Voltage Controlled Magnetic Tunnel Junction

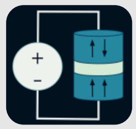


Fig. 1 – VC-MTJ Module
Uses the fixed layer on the bottom as a reference to detect whether the device is parallel, "0" or antiparallel, "1"

## Objectives

- Capability – First demonstration of VC-MTJ Compute-In-Memory (CIM) in large memory arrays
- Power Consumption – Less power will be used in the devices as compared to previous STT-MTJ devices

## Materials

- PCB Design using Altium 2023
- Communication with the PCB using Arduino

## Methods

### Circuit Design

Learning PCB Design starting with no prior experience or knowledge

Based on specifications of what tests are needed to be performed, find components and compare data sheets for:

- Resistance and capacitance of devices
- Bandwidth for high frequencies

### Schematics

Importing the right components into the design software Altium and wiring together each component with net labels

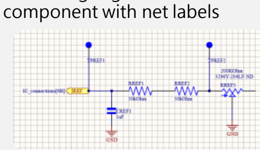


Fig. 2 – Sample Schematic

### Board Layout

Arranging the components onto the PCB layout. Checking to ensure there is proper clearance between components, each connection is valid, and good PCB practices are observed

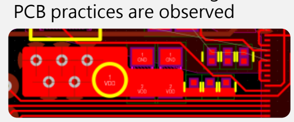


Fig. 3 – Sample PCB Layout

### Testing

After manufacturing, using Python scripts and an Arduino communication system, test the chip for various different operations and limitations of this technology

## Results

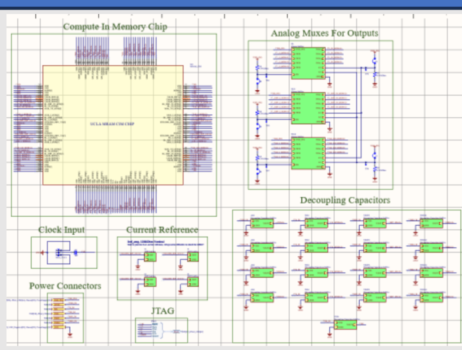


Fig. 4 – Showing a piece of the schematic for the circuit board. This schematic features the main 136 pin MRAM CIM chip in the upper left corner, the analog multiplexers in order to combine the output signals together, and the decoupling capacitors used to filter out the noise that comes from power flowing through the layers of the circuit board.

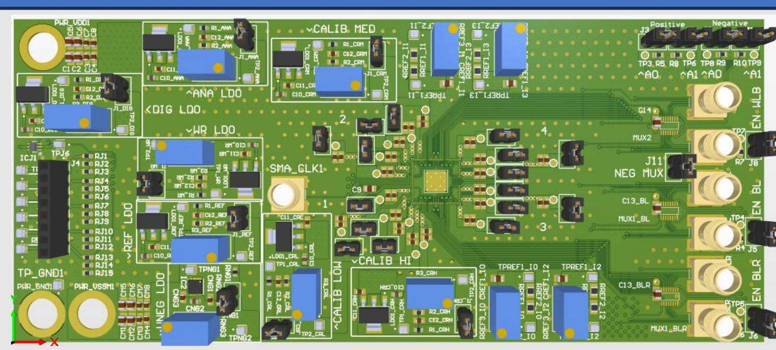


Fig. 5 – The 3D view of the complete PCB board demonstrating how each of the components are laid out and connected together. Within the four layer PCB there is a JTAG interface, LDOs for power regulation, clock inputs, decoupling capacitors, current references, and the analog outputs which can be received through the SMA connectors.

## Conclusion

This PCB will enable the very first demonstration of the promising VC-MTJ device in a compute-in-memory array. Having flexibility in operation, the device will push the limits of technology and redefine computer architecture.

## Discussion

Designing a board of this size from scratch has been a very difficult project and especially with no PCB design experience before, the learning curve has been immense, yet the product is beautiful and many lessons were learned.

## Future Work

Once manufactured, our next steps:

- Interfacing with the IC chip and PCB using JTAG
- Utilizing Arduino for communication, we will run python test scripts to check compute-in-memory operations
- Testing the large memory potential for reading and writing

## References

J. Yang et al., "A Calibration-Free In-Memory True Random Number Generator Using Voltage-Controlled MRAM," ESSDERC 2021 - IEEE 51st European Solid-State Device Research Conference (ESSDERC), Grenoble, France, 2021, pp. 115-118, doi: 10.1109/ESSDERC53440.2021.9631784.

K. Ikegami et al., "Voltage-controlled magnetic tunnel junction based MRAM for replacing high density DRAM circuits corresponding to 2X nm generation," 2017 IEEE International Magnetics Conference (INTERMAG), Dublin, Ireland, 2017, pp. 1-2, doi: 10.1109/INTMAG.2017.8007706.

## Acknowledgements

I would like to thank Professor Pamarti for the opportunity to explore this technology. Vinod and Haris who guided me through many struggles this summer, as well as taught me many fundamental skills. Finally, Spectrum X for funding my efforts.

## FACULTY ADVISOR

**Mitchell Spearrin**

## DAILY LAB SUPERVISOR

**Isabelle Sanders**

## DEPARTMENT

**Mechanical and Aerospace Engineering**


# KATE OBERLANDER

Mechanical and Aerospace Engineering
3rd Year, UCLA


# In-situ IR Laser Spectroscopic Analysis of HF Production through PTFE


## ABSTRACT

Rapid detection of the toxicant HF is essential to the safety of emergency personnel combating fires. HF is produced in electric vehicle and battery fires, and in structural fires in modern developments containing synthetic polymers. We leverage the capability of in-situ infrared laser spectroscopy to detect and analyze the thermochemical structure, including HF concentration, resulting from PTFE combustion with oxygen, a previously unexplored area of research. A laser targeting absorption features of HF and H2O, and a laser targeting those of CO, are pitched through the active combustion zone. Measurements are obtained at varied line-of-sight positions and over a range of fuel grain lengths to resolve 2D evolution of gas temperature and speciation. We assume an axisymmetric and quasi-steady state combustion process allowing our results to be Abel transformed into the radial domain.

As predicted by chemical kinetic simulations, we detected consistently high concentrations of HF in consecutive combustion tests, confirming repeatability of our detection method. Furthermore, we determine temperature from multiple carbon monoxide absorbance lines in the burns. Our spatially resolved, granular thermochemical results demonstrate the robustness of the developed HF sensor and can help anchor PTFE combustion chemical kinetic mechanisms to improve predictions of hazardous HF production. Future research will include sensing and analysis of additional combustion-related toxicants, and development of a prototypical real-time detection device for emergency personnel. This technique can also be used to granularly and quantitatively analyze virtually any axisymmetric combustion, including those with important implications for hybrid rocketry.

SUMMER UNDERGRADUATE RESEARCH PROGRAM

# In-Situ IR Laser Spectroscopic Analysis of HF Production through PTFE Combustion

**Kate A. Oberlander**
**Isabelle C. Sanders, Prof. R. Mitchell Spearrin**
*Department of Mechanical & Aerospace Engineering, University of California, Los Angeles*

## Introduction and Background

**HF** (hydrogen fluoride): **highly reactive toxicant** found in modern fires from synthetic polymers and lithium ion batteries
- Firefighters do not have the capabilities to detect this toxin
- Prior research uses $COF_2$ and $CO_2$ to infer HF concentration[1], [2]
- No prior research has quantified the spatially resolved and granulated chemical structure of HF production

**PTFE**: polymer commonly known as Teflon™
- Forms **HF** during combustion in oxygen

**Infrared laser spectroscopy**: used to analyze chemical structure of PTFE hybrid combustion
- Spectrum of IR lasers scanned through a gas
- Transmitted light is detected
- Data is processed to reveal chemical properties
  - Unique spectra for each molecule, with absorption area determined by factors such as chemical concentration

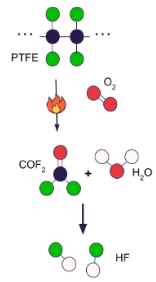


**Figure 1**: Production of HF from PTFE combustion in gaseous oxygen

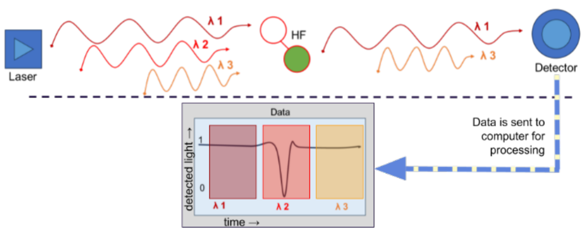


**Figure 2**: IR laser spectroscopy theory, highly simplified

## Objective

We seek to develop a spatially resolved thermochemical analysis of the production of the toxicant HF during hybrid combustions using IR laser spectroscopy

## Materials and Methods

**Composite fuel grains** of PTFE and PMMA in varying lengths
- **PMMA**: polymer commonly known as acrylic
  - Used to induce ignition of PTFE without contributing fluorine

**Graphite gaskets** prevented surface burn propagation

**Gaseous oxygen** flowed through core of fuel
**Nitrogen** used for background measurements and purges.

**IR lasers**: pitched colinearly through active combustion zone at constant velocity of 2.5mm/s

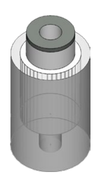


**Figure 3**: Model of PMMA / PTFE composite fuel grain

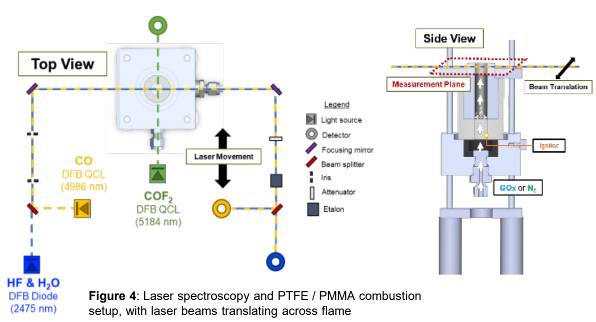


**Figure 4**: Laser spectroscopy and PTFE / PMMA combustion setup, with laser beams translating across flame

**LabView** software: used to gather absorption data, and control laser movement

**MATLAB** software: used to simulate and process absorption data
- MATLAB program was developed to process raw data using the **Beer-Lambert Law**

$$\alpha = -\ln\left(\frac{I_t}{I_0}\right)$$

**Beer-Lambert Law**

**HITRAN** database: used to identify absorption features and associated gas properties of HF, CO and $COF_2$

## Results

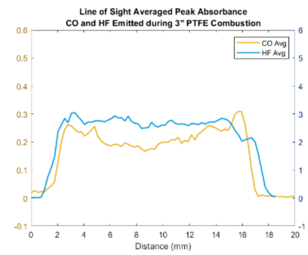


**Figure 5**: Line of sight measurements of peak absorption for 3" PTFE fuel grain

Data successfully collected across 20mm scan of flame
- Burns consistently observed to be **axisymmetric** (*Figure 5*)
  - Abel-transform valid
- No interface burn or surface burn observed
- Averages based on 0.25mm spatial resolution

**High concentrations of HF** detected in each consecutive PTFE burn
- HF absorption area is **maximum** near edge of flame **nearest to fuel grain** (*Figure 7)*
- Implies that HF is primarily produced in a **fuel-rich environment**

**Chemical structure** can be determined from absorption through tomography
- **2236K** average temperature across flame used to determine **concentration of HF**

Peak of HF **broadens** and **moves toward core** as length increases
- **Boundary layer** thickness increases
- **Radial diffusion** of HF from **reaction layer growth** and **turbulent mixing**[3]
- 2in grain has **higher temperature** with proximity to PMMA

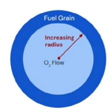


**Figure 6**: Radial cross section of fuel grain and coordinate system of *Figure 7*

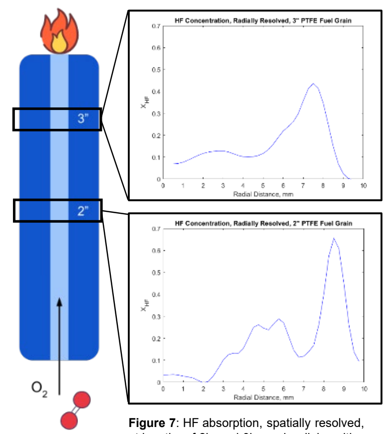


**Figure 7**: HF absorption, spatially resolved, at lengths of 2in and 3in, and radial positions 0mm through 10mm

## Discussion and Conclusion

In-situ laser spectroscopy method **can be used to detect the toxicant HF** in an active combustion with high confidence
- Successfully constructed **two dimensional model** of the hybrid combustion of PTFE and oxygen
- Found that HF is produced primarily in **fuel-rich** environments
- Demonstrated **radial diffusion** of HF in two-dimensional model

HF detection hardware is **increasingly necessary** as emergency personnel encounter dangerous levels of HF during **modern combustion events**:
- Structural fires in modern developments with **synthetic polymers**
- Battery fires involving **electric vehicles**

Future research:
- **Transportable, accessible, fire-resistant** hardware to **help emergency personnel** efficiently and accurately detect HF
- Axisymmetric combustion analysis of virtually **any fuel and oxidizer** combination
  - Important implications for **hybrid rocketry**, a promising field for **space exploration** that is otherwise lacking in granular, spatially resolved analysis

## Acknowledgement and References

This work was supported by the Federal Emergency Management Agency (**FEMA**) and the Summer Undergraduate Research Program (**SURP**). The author thanks SURP leaders **Will Herrera**, **Samanta Negrete**, **Zara Khan** and **Amirhossein Ahmadian** for their outstanding guidance and support. Additionally, the author thanks the Laser Spectroscopy & Gas Dynamics Laboratory, specifically **Prof. Mitchell Spearrin** and **Daily Lab Supervisor Isabelle Sanders**, as well as **Eric Fritz**, **Griffin Houston**, **Barathan Jeevaretanam**, **Erik Juharyan**, **Alex Keller**, **Nick Kuenning**, **Ariya Olaee**, **Kevin Schwarm** and **Yi Yan,** for their excellent insights and collaboration throughout the summer.

[1] Baker, Bertsil B., and Daniel J. Kasprzak. "Thermal Degradation of Commercial Fluoropolymers in Air." Polymer Degradation and Stability, vol. 42, no. 2, 1993, pp. 181–188, https://doi.org/10.1016/0141-3910(93)90111-u.

[2] Miser, Craig S., and W. Randolph Davis. "MEASUREMENT OF CARBONYL FLUORIDE, HYDROGEN FLUORIDE, AND OTHER COMBUSTION BYPRODUCTS DURING FIRE SUPPRESSION TESTING BY FOURIER TRANSFORM INFRARED SPECTROSCOPY." HalonOptions Technical Working Conference, 12 May 1998, pp. 190–203.

[3] Sanders, Isabelle C., et al. "Spatially-resolved characteristic velocity (C*) measurements for hybrid rocket combustion analysis using laser spectroscopy." AIAA SCITECH 2022 Forum, 2022, https://doi.org/10.2514/6.2022-2233.

Flame and firefighting graphics from flaticon.com


## FACULTY ADVISOR

**C.K. Ken Yang**

## DAILY LAB SUPERVISOR

**Julia Bi**

## DEPARTMENT

**Electrical and Computer Engineering**

## LARA SMARANDOIU

Electrical and Computer Engineering
1st Year, UCLA


# Using Mobile Edge Compute to Create an Augmented Reality Application


## ABSTRACT

As mobile applications become more advanced, they require increasing amounts of computing power that is often infeasible for everyday devices to handle. Mobile Edge Compute (MEC) offers a solution to this problem. Computing tasks are offloaded from inefficient devices with limited battery life and processing power to external servers so that applications can have smoother performance, and real-time response rates. The MECA laboratory aims to take advantage of the MEC technology and develop useful augmented reality (AR) applications. MECAL's CloudAR App is an AR application in which users are geographically positioned into a game world and an external server overlays users' camera feed with AR objects based on physical location measurements.

The app is built using Unity, a cross-platform game engine. Unity's WebRTC library provides peer-to-peer (P2P) connection between the client and server for real-time communication and high-speed updates to the app environment. By sending physical GPS coordinates and accelerometer readings to the external server that renders video stream overlays, the client offloads tasks. The CloudAR App is thus able to create an interactable AR environment where users can spawn spheres, squares, and signs with text for all to see in workable areas with radii as large as 500 meters. The MECA laboratory runs the high-speed server that is able to render, on average, 708.18 frames per second with a video streaming and data channel round trip latency of 28.67 ms when one client uses the server. Increasing the number of clients to two puts demand on the server, with round trip latency staying almost consistent but FPS decreasing to 177.08. In this way, the CloudAR App demonstrates the applicability of MEC to previously thought impractical mobile device applications such as AR.

# Using Mobile Edge Compute To Create An Augmented Reality Application

Lara Smarandoiu, Julia Bi, Prof. C.K. Ken Yang
Department of Electrical and Computer Engineering, University of California – Los Angeles

UCLA Samueli School of Engineering
SUMMER UNDERGRADUATE RESEARCH PROGRAM

## Introduction

In this project, the MECA laboratory applies the technology of Mobile Edge Compute (MEC) to an augmented reality (AR) mobile device application, the CloudAR App. In order to extend the capabilities of AR, MEC is used to offload computing tasks to an external edge server from devices that cannot sustain AR's demand for high amounts of processing power. In this way, the CloudAR App is able to provide users with high-quality, low-latency graphics on their devices through the form of useful AR applications.

## Background

As developers create processing-heavy applications, mobile devices' limitations have become prevalent.

- Energy consumption, lack of processing power, and unfavorable high-latency user experiences necessitate technology such as MEC to build applications such as the CloudAR App.
- AR in particular requires a real-time experience, and thus low latency, to be executed well.

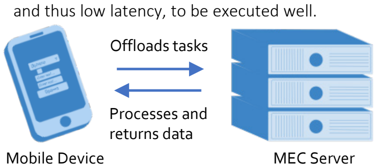



Figure 1: Mobile Edge Compute (MEC)

- Using MEC, a mobile device offloads most, if not all, computing tasks to an external server, allowing for a better user experience through low-latency, higher-quality graphics, and improved efficiency of energy consumption.

## Materials and Methods

- The CloudAR App is built in Unity, a cross-platform game engine, that creates environments using 3D models.
- Unity's WebRTC library establishes a peer-to-peer connection between the client and server, enabling real-time communication. WebRTC allows devices to communicate directly, and is necessary for AR's real-time nature.
- Lastly, an embedded WebSocket server facilities signaling between the server and client.

Figure 2: CloudAR App

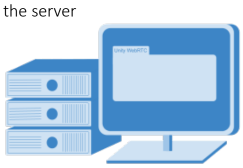


Figure 3: MEC Server and Unity Game Engine

- The ARKit XR Unity Plugin incorporates extended reality (XR) features. Additional tools include Blender Benchmark, a platform used to collect data such as connection latency.

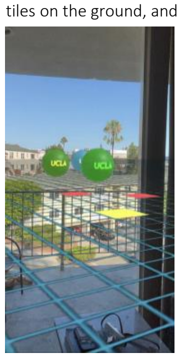
## Geotagging Application

- The MEC server uses the user's real-world location to create a 3D environment users can interact with.
- The client sends the user's GPS coordinates and accelerometer positional information to the server, which places the user in the Unity game world and renders a video stream to overlay the user's camera.
- The server renders an updated video stream when GPS location/accelerometer position changes.

Figure 4: Overlaid Environment Viewed In Different Angles
An example of the geotagging application as the MEC server adjusts the video stream overlay.

## Object Spawning Application

The CloudAR App's 3D environment allows for interaction by spawning objects including spheres, tiles on the ground, and signs with text.



- Users may input custom messages to write on signs, as well as input a specific distance and, on the second page of options, object color.
- A limited number of objects may spawn, and they despawn automatically after a specific time limit.

Figure 5: Spawned Objects, spawned in a variety of colors and at different distances from the user.

## Complete Campus Integration

Through the CloudAR App's interactable environment, MECAL intends to create a community where users on UCLA's campus can communicate through AR. This is made possible by:

- A sandbox-like environment where users can spawn a variety of objects. In the future, users will be able to upload their own objects.
- A continuous floor grid to represent the game world's floor, updating as the user walks.
- A workable 500 meter radius area encompassing the entirety of UCLA's campus, centered at the Royce Hall quad.

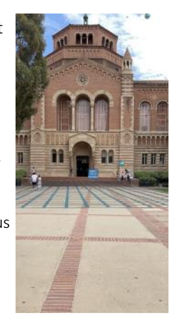


Figure 6: Floor Grid Overlay with spawned objects in front of Powell Library. An example of real-world interaction.

## Server Performance

- When exiting the boundary of the supported 500 meter radius circle that is the "workable area," all overlaid 3D models disappear. When re-entering, objects reappear but may shift and no longer be accurately placed.
- FPS decreases noticeably when nearing the border.


Figure 7: User Approaches Workable Area Border
Represented by the white line, the border represents the limit of the CloudAR App's physical functionality.

## Results

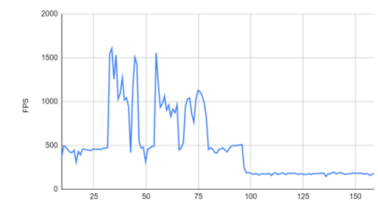



Figure 8: Plot of FPS of the Linux server run by the MECA lab and accessed through the UCLA VPN. 0 to 90 seconds, one client accesses the server. 90 seconds onward two clients access the server. The Linux server receives 708.18 frames on average supporting 1 client and 177.08 frames supporting 2 clients, with an overall average of 488.66 frames.

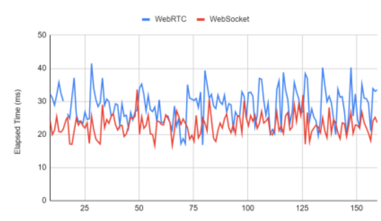



Figure 9: Plot of round trip latency for WebRTC and WebSocket connection on the Linux MECA lab server. The WebRTC server manages video streaming and the data channel, while the WebSocket server facilitates signaling. The WebRTC server has an overall average of 28.67 ms round trip latency, and the WebSocket an average of 22.88 ms. Despite a second client accessing the MECA server around 90 seconds, round trip latency does not change significantly.

## Acknowledgements

I would like to thank Professor Yang and Julia Bi for their invaluable efforts and advice in conducting this project, as well as the Summer Undergraduate Research Program for providing incredible support and, most importantly, the opportunity.


## FACULTY ADVISOR

**Richard Wesel**

## DAILY LAB SUPERVISOR

**Brendan Towell and Beryl Sui**

## DEPARTMENT

**Electrical and Computer Engineering**

## LAURA HUANG

Electrical and Computer Engineering
2nd Year, UCLA


# Inner Tail-biting Convolutional Code (TBCC) with Expurgating Linear Functions (ELFs) as the Outer Code


## ABSTRACT

Communication channels are imperfect due to noise interfering with the transmission, resulting in distorted messages. Ensuring reliable communication systems that minimize the Codeword Error Rate (CER) is important in such scenarios. One effective method to combat this issue is the use of a serially concatenated code with a Cyclic Redundancy Check (CRC) as the outer code and a Tail-biting Convolutional Code (TBCC) as the inner code. This concatenated code leverages the strength of different error-detecting codes to improve overall performance. An Expurgating Linear Function (ELF) is a generalization of the CRC that doesn't restrict the outer code to be cyclic.

For a variety of interesting cases, there are no cyclic codes available so the ELF generalization provides an important insight. This project focuses on the specific case of ELFs used as outer codes for a TBCC and seeks to understand how cyclic codes perform within the larger space of ELFs. By their nature, cyclic codes used for the expurgation of TBCCs will remove or retain all cyclic shifts of a codeword. We proceed to examine the full set of ELFs for cases where the TBCC and ELF redundancy are fixed. Performance is evaluated using union bounds on CER. Our results suggest that if cyclic codes exist in the set of possible ELFs when the TBCC and ELF redundancy are fixed, then the best ELF is a cyclic code. We haven't found a counterexample to this conjecture, which significantly reduces the search space by restricting attention to cyclic ELFs when available.

# Concatenated Tail-Biting Convolutional Codes (TBCCs) with Expurgating Linear Functions (ELFs)

UCLA Samueli School of Engineering
SUMMER UNDERGRADUATE RESEARCH PROGRAM

Laura Huang, [1]Brendan Towell, [1]Beryl Sui, [2]Richard Wesel,
[1]Daily Lab Supervisor, [2]Principal Investigator

## Introduction

**Problem:** Communication systems must overcome noise interference.

- Encoding the message with redundancy lowers the Codeword Error Rate (CER).

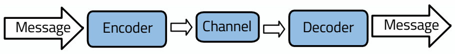



*Fig. 1. Block diagram of a communication system with the message inputted into the encoder, through a noisy channel, and transmitted with the decoder.*

- The Random Coding Union (RCU) bound is an upper bound on the lowest CER that is achievable.
- The Meta-converse (MCU) bound is a lower bound on the lowest CER. The MCU bound itself may not be achievable.

**Goal:** Find codes that approach these bounds more closely

## Background

### Concatenated Code

- A **serially concatenated code** cascades two or more error-detecting / correcting codes to leverage the strength of each individual code and improve overall performance
- Expurgating Linear Functions (ELFs), Cyclic Redundancy Checks (CRCs), and Tail-biting Convolutional Code (TBCC) are examples of error-detecting / correcting codes

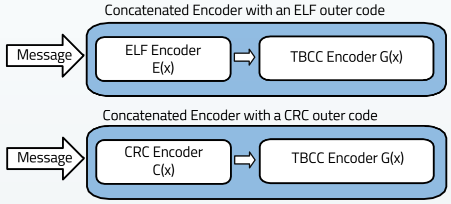



*Fig. 2. Block diagram of a concatenated code with an ELF E(x) outer code and a TBCC G(x) inner code (top). A different construction of concatenated code with a CRC C(x) and a TBCC G(x) is also shown for comparison (bottom). For both diagrams, the outer encoder produces a codeword of length K and is processed by the inner encoder to produce a length N codeword for the channel.*

**Past Research:** Both ELFs and CRCs can be used to expurgate a TBCC. The outer ELF or CRC guides list decoding by checking whether a valid codeword has been correctly received or if the list decoder needs to continue down the list.[1, 3]

### CRCs (and more generally, ELFs)

**Cyclic Codes:**

- Linear codes for which every cyclic shift is also a codeword
  - **Cyclic code** have a generator polynomial that is factor of $x^K - 1$.
  - CRCs are cyclic codes that are traditionally used for error-detection.
  - **Non-cyclic code** don't have a polynomial that is a factor of $x^K - 1$.

**CRC-aided Expurgation of Convolutional Codes (CCs):**

- Lou *et al.*[2] showed that a CRC / ELF designed to expurgate a zero-terminated convolutional code (ZTCC) improved error-detection.
- Sui *et al.*[1] & Wesel *et al.*[3] show that list-decoding of a CRC / ELF concatenated with a TBCC closely approaches the RCU.

**Motivations:**

- The set of ELFs, which doesn't restrict the outer code to be cyclic, contains the set of CRCs as a subset.
- **We seek to understand where the subset of cyclic ELFs (aka CRCs) performs in the space of all possible ELFs.**
- Because the inner TBCC code is cyclic, we expect some CRCs to outperform non-cyclic ELFs. For some values of K, there is no CRC.

**Strategy:**

- Utilize Wesel *et al*, which showed that the ELF - TBCC code can approach the RCU bound, code written by A. Antonini to plot Distance Spectrum Union (DSU) bounds and code written H. Grissett to create the Expurgated Distance Spectrum.[3]

## Materials

MATLAB generated Fig. 3 / 4. by plotting the DSU bounds for various ELFs. C++ code generated Table 1 with the list decoding sieve.

## Methods

**01 PLOT PERFORMANCE OF ELFs**

Plot Distance Spectrum Union (DSU) bounds for all ELFs of degree m for a (561, 753) TBCC with Eq. 1. using Eq.2.

$$P_{cw} \le Q\left(\sqrt{\frac{d_{min}E_s}{\sigma^2}}\right) e^{\frac{d_{min}E_s}{2\sigma^2}} A\left(\frac{-E_s}{2\sigma^2}\right) \quad \text{Eq. 1.}$$

$$A(W) = \sum_{i=0}^{2^v-1} e_i T(W)^K e_i^T - 1 \quad \text{Eq. 2.}$$

where T(W) is the transition matrix for a concatenated convolutional code with an ELF.

**02 ORDER PERFORMANCE OF ELFs**

Determine the design that most minimizes CER (the TBCC with either a cyclic ELF or non-cyclic ELF), using the RCU and MCU as a point of reference. A well-designed ELF expurgated with a TBCC code will approach the RCU / MCU bounds at a certain value of CER. Make sure to order the performance of the ELFs from the worst performing ELF to the best performing ELF.

**03 CREATE THE EXPURGATED DISTANCE SPECTRUM**

Create the Expurgated Distance Spectrum detailing ELF performance with the list decoding sieve. Sort with increasing order of $d_{min}$ and decreasing number of nearest neighbors. Good codes tend to have a high $d_{min}$ and a lower number of nearest neighbors.

## Results

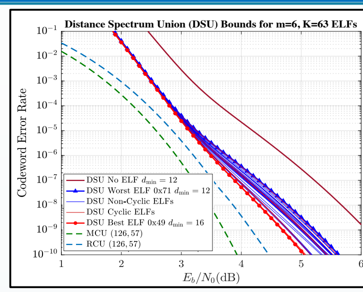



*Fig. 3. Distance Spectrum Union (DSU) bounds for the v=8 TBCC with K=63 bits (and block length N=126) with no ELF (darkRed) and for all possible degree m=6 ELFs. The 20 non-cyclic ELFs are shown in blue and the 12 cyclic ELFs are shown in red. The (126, 57) MCU (dashed green) and RCU bounds (dashed blue) are shown for reference.*

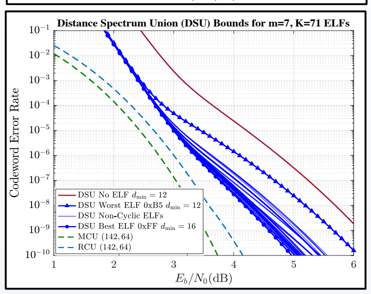



*Fig. 4. Distance Spectrum Union (DSU) bounds for the v=8 TBCC with K=71 message bits (and block length N=142) with no ELF (darkRed) and for all possible degree m=7 ELFs. The 64 non-cyclic ELFs are shown in blue and there are no cyclic ELFs. The (142, 64) MCU (dashed green) and RCU bounds (dashed blue) are also are shown for reference.*

Table 1: Expurgated Distance Spectrum for $m = 6$, $w \le 20$

| E(x) | Cyclic | $A_{12}$ | $A_{14}$ | $A_{16}$ | $A_{20}$ |
|---|---|---|---|---|---|
| 0x71 | No | 1 | 172 | 258 | 9124 |
| 0x7b | No | 1 | 113 | 266 | 8793 |
| 0x59 | No | 1 | 66 | 509 | 10480 |
| 0x6f | No | 1 | 64 | 316 | 9623 |
| 0x47 | No | 1 | 60 | 146 | 9752 |
| 0x51 | No | 1 | 60 | 158 | 10103 |
| 0x4d | No | 1 | 9 | 529 | 9207 |
| 0x4b | No | 1 | 1 | 323 | 9558 |
| 0x55 | No | 1 | 1 | 160 | 9821 |
| 0x45 | No | 2 | 0 | 390 | 8844 |
| 0x5b | Yes | 0 | 189 | 252 | 8505 |
| 0x4f | No | 0 | 168 | 237 | 10124 |
| 0x57 | Yes | 0 | 126 | 126 | 8820 |
| 0x69 | No | 0 | 113 | 234 | 10569 |
| 0x7d | No | 0 | 112 | 310 | 9241 |
| 0x5d | No | 0 | 106 | 357 | 9072 |
| 0x43 | Yes | 0 | 63 | 441 | 10710 |
| 0x67 | Yes | 0 | 63 | 315 | 10710 |
| 0x6d | Yes | 0 | 63 | 252 | 10017 |
| 0x61 | Yes | 0 | 63 | 252 | 8946 |
| 0x73 | Yes | 0 | 63 | 189 | 9513 |
| 0x53 | Yes | 0 | 63 | 126 | 9828 |
| 0x77 | No | 0 | 61 | 377 | 9774 |
| 0x63 | No | 0 | 56 | 156 | 8990 |
| 0x5f | No | 0 | 55 | 223 | 9485 |
| 0x41 | No | 0 | 5 | 302 | 10048 |
| 0x79 | No | 0 | 5 | 211 | 10163 |
| 0x6b | No | 0 | 1 | 376 | 10012 |
| 0x65 | Yes | 0 | 0 | 378 | 9513 |
| 0x75 | Yes | 0 | 0 | 378 | 9198 |
| 0x7f | Yes | 0 | 0 | 189 | 9954 |
| 0x49 | Yes | 0 | 0 | 126 | 9135 |

*Table 1. Expurgated Distance Spectrum for degree m = 6 ELFs for the v=8 (N = 126, K=63) TBCC (561, 753). $A_\#$ is the number of neighbors at distance #. Note that for cyclic ELFs $A_\#$ is always a multiple of K = 63.*

## Discussion

From Fig.3:

- Table 1 shows that **the best 4 ELFs are cyclic ELFs** whereas **the worst 10 ELFs are non-cyclic ELFs**.
  - In the middle of the table, both cyclic ELFs and non-cyclic ELFs appear
- Using the MCU as a reference, it's not surprising to see that the DSU's lie above the MCU lower bound for lower Eb/No and converges to the bound for higher Eb/No at a certain CER.

From Fig.4:

- Non-cyclic ELF 0xFF is the best ELF
  - Interesting since we know that the all-ones is generally cyclic for certain cases of K (not K = 71).

## Conclusion/ Future Work

**Findings:**

- For the K=63 case, where cyclic ELFs are available, the best ELF was a cyclic ELF.
  - We conjecture that when a cyclic ELF is available, then **the best ELF will always be a cyclic ELF**
- For the K=71 case, where no cyclic ELFs are available, the method of checking all ELFs is required
  - We conjecture that 0xFF was the best since it was almost cyclic, may potentially be cyclic for K = 70 or K = 72 and was only one off – we can't say anything with certainty.

**Future:**

- See if it is always true that when a cyclic ELF is available, then the best ELF will always be a cyclic ELF
- See if 0xFF is cyclic for a K close to K = 71 and see if a cyclic ELF can exist that way

## References

[1] W. Sui, H. Yang, B. Towell, A. Asmani, and R. D. Wesel, "High-rate convolutional codes with CRC-aided list decoding for short blocklengths," in ICC 2022 - IEEE International Conference on Communications, 2022, pp. 98–103.
[2] C.-Y. Lou, B. Daneshrad, and R. D. Wesel, "Convolutional-code-specific CRC code design," IEEE Trans. on Communications, vol. 63, no. 10, pp. 3459–3470, 2015.
[3] R.D.Wesel, A.Antonini, L.Wang, W.Sui, B.Towell, and H. Grissett,"ELF codes: concatenated codes with Expurgating Linear Function as the Outer Code" [Online]. Available: https://arxiv.org/pdf/2306.07467.pdf

## Acknowledgements

I'd like to thank Dean Wesel for his guidance while I was with the Communication System Lab at UCLA. I especially thank Brendan and Beryl for their persistent insight and assistance with technical issues. I'm also grateful for UCLA SURP for making this poster possible. This work was supported by the National Science Foundation (NSF) under Grant CCF-2008918.


FACULTY ADVISOR

Danijela Cabric

DAILY LAB SUPERVISOR

Tianyi Zhao

DEPARTMENT

Electrical and Computer Engineering

MARVIN MOK

Electrical and Computer Engineering
2nd Year, UCLA


## Identification of Bluetooth Devices via RF Finger-Printing

### ABSTRACT


Radio frequency fingerprinting (RFF) is a physical layer identification technique that can be used for various purposes such as security and device tracking. Due to imperfections in hardware, each device will have its own distinctive "fingerprint" when transmitting signals. Using machine learning models such as support vector machines or convolutional neural networks, individual devices can be identified via their own fingerprints. Bluetooth signals are of interest to fingerprint. Bluetooth is a wireless communication used for low range and low power connectivity between devices. Specifically, BLE employs a modification of a technique called frequency hopping spread spectrum (FHSS). Frequency hopping is when a transmitter changes its carrier frequency across multiple channels in a large spectral band, adding reliability to a connection.

First, transient-based features such as the Shannon entropy and transient energy were extracted from highly oversampled Bluetooth signals from 27 different phone makes-and-models. Then, these features were used in machine learning models to correctly identify different phones. Next, frequency hopping, Bluetooth Low Energy (BLE) like signals were generated in MATLAB and transmitted over the air using 10 different ADALM-PLUTO software defined radios (SDR) to create a dataset of undersampled signals. Steady-state features such as carrier frequency offset and in-phase quadrature offset were extracted, and machine learning models were then used to identify each radio through its fingerprint. Further research includes reducing the bandwidth of the receiver radio to only analyze specific channels.

# Identification of Bluetooth Devices via RF Fingerprinting

Marvin Mok, Tianyi Zhao, Prof. Danijela Cabric
Cognitive Reconfigurable Embedded Systems (CORES) Lab
Department of Electrical and Computer Engineer | University of California, Los Angeles

UCLA Samueli School of Engineering

SUMMER UNDERGRADUATE RESEARCH PROGRAM

## Introduction

Radio frequency fingerprinting (RFF) is a physical layer identification technique that can be used for various purposes such as security and device tracking. Due to imperfections in hardware, each device will have its own distinctive "fingerprint" when transmitting signals. Using machine learning models such as support vector machines, Bluetooth devices were identified via their own fingerprints. Then, a dataset was created to classify frequency hopping Bluetooth Low Energy Devices signals.

## Background

Bluetooth signals are of interest to fingerprint. Bluetooth is a wireless communication used for low range and low power connectivity between devices. Bluetooth Bluetooth use the Gaussian frequency shifting key (GFSK) modulation scheme to encode signals. GFSK shifts frequency up or down from the center frequency a frequency deviation to encode a 1 or 0. Furthermore employ a modification of a technique called frequency hopping spread spectrum (FHSS). Frequency hopping is when a transmitter changes its carrier frequency across multiple channels in a large spectral band, adding reliability to a connection. Both Bluetooth and Bluetooth Low Energy have a standard hop rate of 1500 hops per second, with 80 different 1MHz bandwidth channels for Bluetooth and 40 different 2MHz bandwidth channels for BLE.

For ideal Bluetooth and BLE signals, the constellation diagram of the in-phase and quadrature data would be a perfect circle. However, due to hardware imperfections, the signal becomes warped with equation 1. Carrier frequency offset (CFO) applies a time-dependent phase shift, I/Q offset moves the center of the constellation, and I/Q imbalance warps the circle in an ellipse.

$$y'(t) = A \times \left[(1-\frac{\epsilon}{2})cos(\omega(t)t - \frac{\phi}{2}) + I + \right.$$
$$\left. j\left((1+\frac{\epsilon}{2})sin(\omega(t)t + \frac{\phi}{2}) + Q\right)\right] \times e^{j(\phi_o + 2\pi f_o t)}$$

Equation 1: An imperfect Bluetooth signal, where $A$, $\frac{1+\epsilon}{1-\epsilon}$, $\omega(t)$, $\phi, I, Q, \phi_o, f_o$ are the normalized amplitude, I/Q imbalance, carrier frequency, I/Q phase imbalance, in-phase offset, quadrature offset, phase offset and carrier frequency offset, respectively.

## Data Collection

Two Datasets were used to classify different Bluetooth and BLE signals.

The first dataset [1] consists of Bluetooth signals from 27 different off-the-shelf make-and-models of phones. Additionally, two phones for each model were used. 150 captures per phone (voltage-time), each ~10 µs long, were collected in a laboratory environment using a commercial off-the-shelf antenna and oscilloscope.

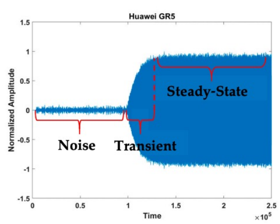



Figure 1: Sample Bluetooth signal split into its three parts.

A second dataset was created to analyze frequency hopping BLE signals. BLE signals were generated in MATLAB and transmitted through a ADALM PLUTO software defined radio (SDR). Since the PLUTO radio only has a bandwidth of 20MHz, signals only hopped between 7 channels. 8 different radios were used for transmission and 1 for receiving, with a separation of ~30 cm. Raw I and Q samples were collected on the 5th floor of Engineering IV at UCLA at a 40 MHz sample rate. 70 identical BLE packets were transmitted over the air in a pseudo random hopping sequence per capture, with 150 captures per transmitter. Packets were then separated using a spectrogram and labelled with their channel and transmitter number for a total of 84,000 packets.

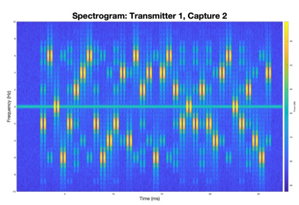



Figure 2: Sample Spectrogram for frequency hopping signals. The bright bands indicates packets hopping between different channels

## Feature Extraction

Transients of signals were detected using equations 2 and 3. 16 features were then extracted using Hilbert and Hilbert Huang Transforms to get I/Q samples and Time-Frequency Energy Distributions:

1-9. Standard Deviation (STD), skewness and Kurtosis of Instantaneous phase, amplitude and Frequency
10. Transient duration
11. Total transient energy
12. Shannon Entropy of phase
13. 3rd order polynomial fitting of energy-time distribution
14-16. STD, skewness and kurtosis of energy-time distribution

For the second dataset, steady-state features were extracted to analyze frequency hopping signals. First, carrier frequency offset (CFO) was estimated using the 8 symbol BLE preamble. An ellipse fitting algorithm was used on the signal constellation to estimate I/Q offset. Then, CFO and I/Q offset were jointly estimated using Nesterov Gradient Descent with the following cost function,

$$min_{f_o,\phi_o,A,\epsilon,\phi,I,Q} F = ||y' - y||^2 =$$
$$|Real\{y'\} - Real\{y\}|^2 + |Imag\{y'\} - Imag\{y\}|^2$$

where y is the ideal Bluetooth signal and y' is equation 1.

$$V(x) = moving\_variance(unwrapped\_phase(signal))$$

$$VT(x) = |V(x+1) - V(x)|$$

$$VT(x), VT(x+1)...VT(x+W) > T, W = 20$$

$$T = 0.2 * (max(VT) + mean(VT(1:100))$$

Equation 2: Phase-based transient start detection

$$Pks(x) = peaks(moving\_avg(amplitude(signal))))$$

$$\text{Leftmost Peak Satisfying:} \quad Pks(x) > T$$
$$T = (min(Pks)) + mean(Pks(end - W : end))) * 0.85, W = 5$$

Equation 3:Amplitude-based transient end detection equation:

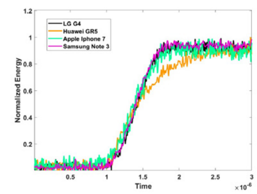



Figure 3: Sample signals overlayed. While the steady-state looks very similar, the transient are more clearly differentiable.

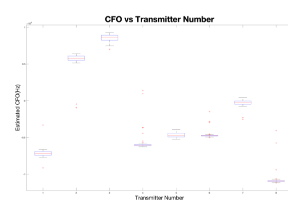



Figure 4: Boxplot of CFO for different SDRs. Transmitters are nearly separable just by their CFOs, with exceptions of transmitters 5 and 6, which have similar fingerprints.

## Classification and Results

| Sample Rate | Unsmoothed | Smoothed |
|---|---|---|
| 250 Msps | 89.31% | 92.51% |
| 10 Gsps | 95.74% | 97.58% |

Table 1: Classification performances of SVM for different sample rates

| ML Model | Unsmoothed | Smoothed |
|---|---|---|
| LDA | 87.09% | 89.31% |
| RBF-SVM | 89.18% | 93.08% |
| Random Forest | 92.35% | 96.46% |

Table 2: Classification performances of different machine learning models for frequency hopping BLE signals.

Classification was done in Python using scikit-learn libraries.

Two sampling rates, one highly oversampled and one under sampled, were used to classify Bluetooth signals between different phone models using transient-based fingerprinting. Features were fed into a support vector machine (SVM) using a radial basis function (RBF) kernel.

Three different classical machine learning models were used to test their viability to identify frequeny hopping BLE signals between individual ADALM PLUTO SDRs: Linear Discriminant Analysis (LDA), SVM with RBF kernel, and a Random Forest Classifier.

Data was split 48-40-12 for training, testing, and validation to find cross-validated performance accuracy. Furthermore, feature smoothing using a fourth-ordered median filter was used to improve classification performance.

## Conclusions

Bluetooth sand BLE signals were identified using their RF fingerprints with high accuracy using classical machine learning models, both for transient-based and steady-state fingerprinting. Frequency hopping did not seem to affect the fingerprint significantly. For steady-state fingerprinting, some transmitters fingerprints were too similar to consistently identify.

Further steps include transient-based fingerprinting for the frequency hopping dataset and comparing deep learning models such as convolutional neural networks to classical ML methods.

## References

[1] Uzundurukan, E.; Dalveren, Y.; Kara, A. A Database for the Radio Frequency Fingerprinting of Bluetooth Devices. Data 2020, 5, 55.
[2] Ali, Aysha & Uzundurukan, Emre & Kara, Ali. (2017). Improvements on transient signal detection for RF fingerprinting. 1-4. 10.1109/SIU.2017.7960417.
[3] Ali, Aysha & Uzundurukan, Emre & Kara, Ali. (2019). Assessment of Features and Classifiers for Bluetooth RF Fingerprinting. IEEE Access. 7. 50524-50535. 10.1109/ACCESS.2019.2911452.
[4] H. Givehchian *et al.*, "Evaluating Physical-Layer BLE Location Tracking Attacks on Mobile Devices," *2022 IEEE Symposium on Security and Privacy (SP)*, San Francisco, CA, USA, 2022, pp. 1690-1704, doi: 10.1109/SP46214.2022.9833758.

## Acknowledgements

I would like to thank Prof .Danijela Cabric, Tianyi Zhao, and the rest of the CORES lab for their help and guidance. I would also like to thank SpectrumX and the Summer Undergraduate Research program for funding this project.


**FACULTY ADVISOR**

**Tony Nowatzki**

**DAILY LAB SUPERVISOR**

**Sihao Liu**

**DEPARTMENT**

**Computer Science**

## MAXIM ZHULIN

Computer Science and Engineering
4th Year, UCLA


# Improving FPGA DSP Usage in OverGen's Functional Units

## ABSTRACT


Field Programmable Gate Arrays (FPGAs) are a powerful alternative to CPUs and custom microchips for running computational workloads. While High Level Synthesis (HLS) is the mainstream programming approach for FPGAs, OverGen is an overlay architecture for FPGAs that has proven to be highly competitive with HLS-based designs. Overgen contains functional units (FUs) implemented with FPGA Digital Signal Processing (DSP) resources, which handle the arithmetic and logic operations of the processing elements. However these FUs are not implemented in the most hardware efficient way - different operations such as multiplication and addition are implemented on separate DSP slices, and their results are multiplexed. This not only increases DSP usage, but also power consumption. The purpose of this research is to fuse the arithmetic and logic operations by using the same DSP slices for different purposes, thus saving in FPGA resources for each FU. The OverGen overlay generator would then be able to map more processing elements onto an FPGA than before, and possibly achieve better performance on workloads.

Furthermore, this research is a stepping stone to look at more generalized ways of mapping desired operations such as multiply-accumulate onto arbitrary DSP networks, to allow the overlay generator to use the DSP slices more efficiently. While the implementation in this research is a custom hardware fusion, perhaps fully exposing the capabilities of the DSP and the network that connects them to the software compiler could yield even better results.

# Improving FPGA DSP Usage in OverGen's Functional Units

UCLA Samueli School of Engineering
SUMMER UNDERGRADUATE RESEARCH PROGRAM

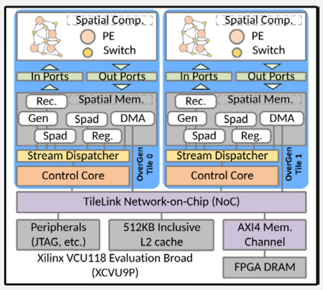
Maxim Zhulin, Sihao Liu, Prof. Tony Nowatzki
Department of Computer Science, University of California Los Angeles

## Background

OverGen is an overlay generator for Field Programmable Gate Arrays (FPGA), which are becoming increasingly common as replacements for CPUs and custom microchips.

In order to do computation, OverGen has many interconnected Processing Elements (PE), which in turn contain functional units (FU) specializing in different math and logic operations such as 64 bit multiply, or floating point divide.

The current implementation of FUs does not efficiently reuse Digital Signal Processing (DSP) hardware on the FPGA.




An example OverGen two tile configuration.

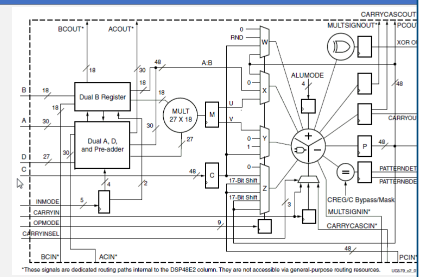
## Objective

Explore ways of fusing operations such as addition, multiplication, and bitwise logic. This could significantly reduce the number of DSP slices used per functional unit, allowing the overlay generator to map more processing elements onto the FPGA, while also reducing power consumption and wasted resources.

## Methods and Materials

While most of the OverGen hardware is written in Chisel, the functional units use Xilinx generated IP, so we will be writing the fused functional units in Verilog.

The figure to the right shows the architecture of the DSP slice, which we will use to implement 64/32/16-bit multiplication, addition, subtraction and bitwise logic.




DSP48E2 Architecture (AMD Xilinx Documentation Portal)

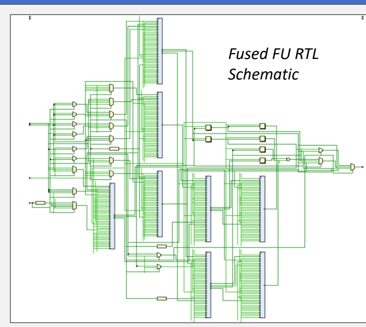
We will use Xilinx's Vivado to synthesize the Verilog code onto an FPGA.

VERILOG VIVADO XILINX

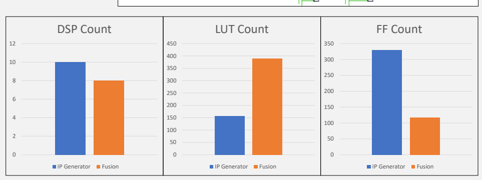
Implement Fused Functional Unit in Verilog → Simulate to Verify Functionality → Synthesize on FPGA to get Resource Utilization

## Results

- Vivado Synthesis outputs 3 relevant hardware statistics: LUT (look up table), FF (flip flop) and DSP (Digital Signal Processing) counts. We synthesized Xilinx IP 64-bit addition and multiplication as a baseline for comparison. The most important statistic to this research is DSP use.
- The results show that there is a **20%** decrease in DSP use, a **65%** reduction in FF use and a **147%** increase in LUT use.
- It is important note that the baseline for comparison used here is beyond a worse case scenario: there are no input or output MUXs for the IP implementation. A fused implementation will likely achieve an overall decrease in LUT use, and even greater decreases in DSP and FF uses in a real scenario. Further testing must be done with full integration into OverGen for true improvement numbers.




| Function | Type | LUT | FF | DSP |
|---|---|---|---|---|
| Add | IP Generator | 123 | 247 | 0 |
| Mult | IP Generator | 34 | 82 | 10 |
| Add+Mult | IP Generator | 157 | 329 | 10 |
| Fused AddMult | Fusion | 389 | 117 | 8 |




## Discussion

- Fusing operations to run on the same DSP slices clearly has an advantage over multiplexed Xilinx IP designs. The synthesis confirms that the fused design indeed uses less DSP slices, while having more flexibility to implement various operations even beyond addition, multiplication, and logic.
- This research was mostly an engineering problem, but it is a stepping stone to having more complicated operation fusion in OverGen's compiler. The goal in OverGen's hardware is to use the FPGA resources as efficiently as possible. Given the DSP savings shown in the results, this project should show an immediate improvement in OverGen's area utilization in PEs, but perhaps the more valuable contribution is the knowledge gained in how DSPs are commonly connected together to perform more complicated and wider operations.
- Further research could investigate delegating the operation fusion to the OverGen software compiler, rather than hardcoded hardware implementations.

## References

Liu, Sihao, et al. "OverGen: Improving FPGA usability through domain-specific overlay generation." *2022 55th IEEE/ACM International Symposium on Microarchitecture (MICRO)*. IEEE, 2022.

Xilinx. "UltraScale Architecture DSP Slice." *Xilinx Documentation Portal*, Aug. 2021, Accessed 22 Aug. 2023.

Weng, Jian, et al. "Dsagen: Synthesizing programmable spatial accelerators." *2020 ACM/IEEE 47th Annual International Symposium on Computer Architecture (ISCA)*. IEEE, 2020.

## Acknowledgements

I would like to thank Prof. Tony Nowatzki and Sihao Liu for their oversight and guidance, as well as for the opportunity to work and learn in their lab.
I would also like to thank SURP for organizing valuable presentations and for the opportunity to publish in a journal.
Finally, I would like to thank NSF for providing the funding for my summer research.


FACULTY ADVISOR
**Rob Candler**

DAILY LAB SUPERVISOR
**Vaibhav Sharma**

DEPARTMENT
**Electrical and Computer Engineering**

**RACHEL YEN**

Electrical and Computer Engineering
1st Year, UCLA


## Low Frequency Compact Magnetic Shielding Using Thick-Film Electroplated Permalloy


### ABSTRACT

As electronic equipment such as atomic clocks, transmission cables, and microprocessor-controlled devices become increasingly miniaturized, the magnetic fields they produce will interact in ways never seen before. Problems such as shifts in atomic transition frequencies within chip-scale atomic devices arise when we compact systems to the micro-scale. My research focuses on using alternating layers of ferromagnetic and diamagnetic material to produce a micro-scale shield that prevents low-frequency magnetic fields from entering the protected area. Permalloy has lower magnetic reluctance due to its higher magnetic permeability and thus transmits the magnetic field through materials more easily, effectively shielding the inner components by redirecting the field lines. My research extends the work of Wu et al., which demonstrated that large numbers of thin shields can provide a much higher shielding factor than a single shield with the same material thickness.

By parameterizing the dimensions of shield components within COMSOL Multiphysics, I ran various simulations that accurately portray how physical shields would behave under predetermined conditions. Specifically, using the “Magnetic Field, No Currents” package, I simulated a magnetic field on shield designs to determine the resulting magnetic flux density on the shield and surrounding areas. Through these simulations, I found that the optimal solution is a five-layer slanted shield with 300μm x 300μm through silicon vias (TSVs) every 200μm. Ultimately, the miniaturization of shielding allows for the isolation of superconducting electronic chips from each other and the external environment, which opens up a multitude of possibilities for microelectronics.

# Low Frequency Compact Magnetic Shielding Using Thick-Film Electroplated Permalloy

UCLA Samueli School of Engineering

SUMMER UNDERGRADUATE RESEARCH PROGRAM

**Rachel Yen, Vaibhav Sharma, Siyuan Liu, Jimmy Wu, Prof. Rob Candler**
Department of Electrical and Computer Engineering, University of California – Los Angeles

## Introduction

While electric fields can be easily shielded by conductors, high permeability material is required to redirect **magnetic fields** around enclosed spaces. With this, Wu et al. proposed that many thin shields can provide a much higher **shielding factor** than a single shield with the same material thickness. In this research, layers of **copper** and **permalloy** are used when simulating the shields.

The flexibility in shield thickness and shape design enables **compact integration** of magnetic devices for microelectronics and atomic, molecular, and optical (AMO) microsystems. The previous design of a three-layer, eighteen 300µm x 300µm through silicon via (TSV) shield produces a shielding factor of 5.718. In my research, I utilized the **Finite Element Modeling (FEM)** under COMSOL Multiphysics to demonstrate the optimization of current shielding designs by parameterizing components to obtain the highest shielding factor within limitations.

## Background and Objectives

**Shielding Factor**

The shielding factor, characterized by the external field strength over the internal field strength, provides a quantitative result to measure how effective the shield is.

**Single Cylindrical Shell**

$$S = \frac{\overrightarrow{B_0}}{\overrightarrow{B_1}} = \frac{\mu_{r1} t_1}{d_1}$$

*d = diameter of shield*

*$B_0$ = external field density*

*$B_1$ = internal field density*

*$\mu_r$ = relative permeability*

*t = thickness of layer*

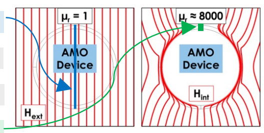


Fig. 1: Magnetic shielding (field line redirection) shown on high and low permeability material. The external magnetic field, $H_{ext}$, and the internal magnetic field, $H_{int}$, are shown.

Magnetic dipoles in the ferromagnetic material align with external magnetic field, inducing a magnetic field that originates from the material, redirecting the external magnetic field.

Layers of copper provide gaps that minimize magnetic interference between successive layers of permalloy.

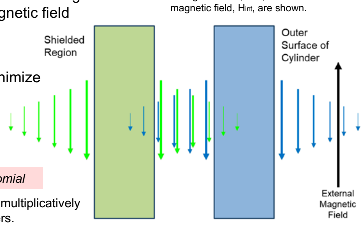


Fig. 2: Two layers of high permeability material, shown in green and blue, in the presence of an external magnetic field. The induced magnetic field is shown in green and blue arrows.

**n Layered Shell**

$$P_0 = 1$$

$$S = \sum_{k=0}^{n} P_k$$

*P = shielding polynomial*

Shielding factor scales multiplicatively with the number of layers.

$$P_1 = \sum_{i=1}^{n} S_i = S_1 + S_2 + \cdots + S_n$$

$$P_2 = \sum_{i=1}^{n-1} \sum_{j>i}^{n} S_i d_{i,j} S_j = S_1 d_{1,2} S_2 + S_1 d_{1,3} S_3 + \cdots + S_2 d_{2,3} S_3 + \cdots + S_{n-1} d_{n-1,n}$$

Using high and low permeability material in successive layers, my goals are to:

1. **Calculate the magnetic shielding factor of the shield given the external magnetic flux density**
2. **Simultaneously push the limits of minimum size, maximum shielding factor, and minimum cost**

## Materials and Methods

### 1 Simulate Shield and Run Simulation

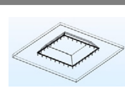


Fig. 3: An example of a shield simulation design.

- Build the shield components in COMSOL – cap layers, through silicon vias (TSV), silicon substrate, and bottom permalloy layer
- Set materials and permeability – permalloy (80% Ni 20% Fe): μ = 1000, copper (100% Cu): μ = 1
- Run the "Magnetic Fields, No Currents" physics interface to model an external magnetic field on the simulated shield and superconducting chip

### 2 Gather Data and Analyze Plots

Google Sheets

- Calculate and plot the shielding factor as a function of position on the shield
- Parameterize variables in COMSOL Multiphysics to find the optimal number of layers, thickness of layers, total number of TSVs, and size of TSVs

### 3 Determine Optimized Shield

- Analyze data for lowest magnetic flux density for each number (n) and size (pw) of layers and TSVs
- Conclude optimal shield for highest shielding factor

COMSOL MULTIPHYSICS

## Results

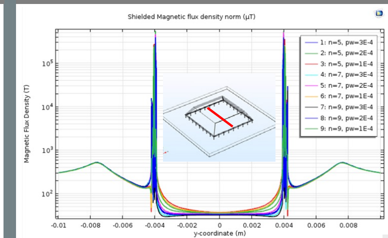


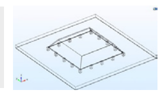
Fig. 4: Shielded Magnetic Flux Density is graphed as a function of position on the shield. The peaks are at the shield side walls, showing the permalloy attracting the magnetic field lines.

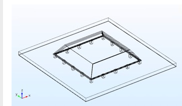
***High Permeability Shielding (μ = 1000) in Background Magnetic Field.***
The magnetic flux density for nine combinations of number of TSVs per side (5, 7, 9) and side length of each TSV (300µm, 200µm, 100µm) is shown. The background magnetic field is 200µT in the horizontal direction.

***n** = number of TSVs per side* ***pw** = size of TSVs*

***Confirming the Direct Relationship between Area of Total Permalloy and Shielding Factor***
Fig. 4 shows the resulting magnetic flux density in the center of the shield. The configurations with more permalloy material have the lowest magnetic flux density, and vise versa.

Before: 3-layer, 18 300µm x 300µm TSVs
**S = 5.718**





After: 5-layer, 18 300µm x 300µm TSVs
**S = 6.048**

Figs. 5 & 6: Shield designs before and after optimizing.

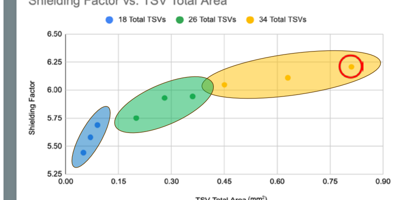
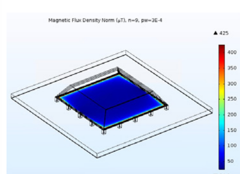


Fig. 7: Shielding factor for various configurations is graphed as a function of total permalloy area. The optimal design found is highlighted.

***Plot of the shielding factor as a function of TSV Total Area***
The trend of shield configurations follows a logarithmic pattern for increasing numbers and sizes of TSVs. As the number and the size of TSVs increase, the shielding factor also increases. Data points with the same number of TSVs follow a linear pattern within the subgroup.

## Conclusion and Discussion




The optimal shielding configuration that produced a maximum shielding is a 5-layer, 34 total 300µm x 300µm TSVs shield. The optimal shielding factor is calculated to be 6.048, a 5.77% increase from the original factor of 5.718. Through the early stages of shields in the sub-millimeter regimes designed, this work demonstrates the use of multilayered shields that enables compact integration of magnetic devices.

$$S = \frac{\overrightarrow{B_0}}{\overrightarrow{B_1}} = \frac{200\ \mu T}{33.069\ \mu T} = 6.048$$

Fig. 8: COMSOL simulation of magnetic flux density within the shield for the optimal shielding configuration is shown.

## Limitations and Future Work

- The **materials** I used to simulate the shield were restricted to permalloy and copper. Other materials, such as Metglas and mumetal, that are expected to produce better shielding due to their higher relative permeability, are currently impractical to use due to fabrication, cost, and ownership reasons. The use of material with **different compositions** can be explored.
- The **shape** of the shield simulated was chosen with fabrication processes kept in mind. The slanted shield shape simulated is due to the buildup of material along the carrier wafer sidewalls that occurs during electroplating. Other shapes can be explored, especially to fit specific chip modules or other devices.
- The concept of shielding can also be expanded to shielding an **internal field** instead of an external field, for example, a current-carrying wire that induces a magnetic field within the shield. Instead of protecting internal devices from external fields, the internal fields would be prevented from spreading externally.

## References

[1] Wu, Jimmy, Ling Li, Jere C. Harrison, and Rob N. Candler. 27 July 2017. "Compact Magnetic Shielding Using Thick- Film Electroplated Permalloy." Doctoral dissertation, UCLA.
[2] Liu, Siyuan, Rob N. Candler. 2023. "Millimeter Scale Magnetic Field Manipulation in Magnetic Resonance Imaging Radio Frequency Coil and Magnetic Shielding Applications." Doctoral dissertation, UCLA.
[3] Sherman E. Peek, Siyuan Liu, Benjamin Pound, Harshil Goyal, George Hughes, Robert Candler, Mark L. Adams, Michael C. Hamilton. "Towards Low Profile Passive Magnetic Shielding for Cryogenic Electronics: Simulation and Material Characterization." GOMACTech 2023, March 2023.

## Acknowledgements

I would like to thank Professor Rob Candler and Vaibhav Sharma for their guidance throughout the summer. I would also like to thank the Summer Undergraduate Research Program (SURP), the Samueli Research Scholars (SRS), and the Sensors and Technology Lab (STL) for organizing and funding my research.


## FACULTY ADVISOR

**Lihua Jin**

## DAILY LAB SUPERVISOR

**Chen Wei**

## DEPARTMENT

**Mechanical and Aerospace Engineering**

# RICHARD ZHOU

Mechanical and Aerospace Engineering
1st Year, UCLA


# Investigating the Fracture Mechanics of Liquid Crystal Elastomers


## ABSTRACT

Liquid Crystal Elastomers (LCEs) are a unique type of soft material that incorporates rod-like liquid crystals (LCs) into a flexible polymer network. The reorientation of LCs gives rise to large spontaneous deformation. As such, LCEs have recently been under study in numerous applications including soft robotics, biomedical devices, and artificial muscles/tissue. As these applications are further developed, it is increasingly important to understand the fracture behavior of LCEs to predict and optimize the lifetime of LCE components. The analysis of the energy release rate in LCEs can be complicated, considering the director rotation and dissipation energy associated with the viscoelasticity of the network and the director.

The intricate interplay of the network and director give rise to a complex fracture phenomenon, where the crack may propagate with an inclined angle resulting in a mixed-mode fracture. To characterize LCE fracture behavior, we first focused on monodomain LCEs with a director parallel to the loading direction under Mode-I loading. We then analyzed the fracture energy under different loading rates using the pure shear test; delayed fracture was explored via the relaxation and creep tests. Our results show that the loading rate has a direct correlation with fracture energy. Furthermore, for delayed fracture, the threshold displacement was approximated and the rate of crack propagation was compared across different holding values. In the future, we will explore behavior in a fatigue test and different loading modes.

# Investigating the Fracture Mechanics of Liquid Crystal Elastomers

UCLA Samueli School of Engineering
SUMMER UNDERGRADUATE RESEARCH PROGRAM

Richard Zhou, Chen Wei [a], Lihua Jin [b]
Department of Mechanical and Aerospace Engineering, University of California – Los Angeles
[a] Daily Lab Supervisor, [b] Principal Investigator

## Motivation

Liquid Crystal Elastomers (LCEs) are a unique type of soft material that incorporates rod-like liquid crystals (LCs) into a flexible polymer network. The reorientation of these LCs can give rise to large spontaneous deformation.

LCEs have recently been under study in numerous applications including soft robotics and biomedical devices.

Therefore, it is increasingly important to understand the fracture behavior of LCEs to predict and optimize the lifetime of LCE components.

## Objective

To experimentally observe the rupture behavior of monodomain LCE samples with a director parallel to the applied load by finding the fracture energy under different loading rates and exploring delayed fracture.

## Background

### LCE Orientation:

In nematic monodomain LCEs, LC mesogens tend to align in a specific orientation; the average direction is referred to as the **director**. The angle between the director and the loading direction is used to describe the relative tilt of the LCE.

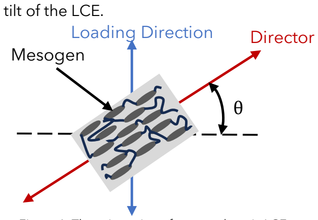



Figure 1. The orientation of a monodomain LCE

### Fracture Energy:

As a crack propagates through a sample, the **total potential energy lost** per unit increase in crack growth area can be quantified as the **energy release rate**. The critical energy release rate at crack propagation is defined as the **fracture energy**.

### Monotonic/Static:

**Monotonic Loading:** Displacement increases at a constant rate. Corresponds to fast fracture.
**Static Loading:** Strain/stress is held constant. Corresponds to delay fracture.

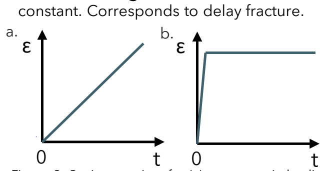



Figure 2. Strain over time for (a) a monotonic loading case and (b) a static loading case

## Materials and Methods

### 1. Fabrication and Setup

Two-stage thiol-acrylate Michael addition-photopolymerization (TAMAP) reaction [1]

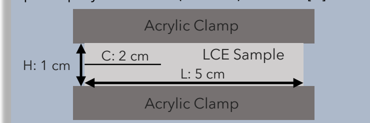



Figure 3. Sample geometry and clamping setup for a pre-cut sample

### 3. Analysis

**Pure Shear Test:**

I. Determine the point of fracture of the pre-cut sample through video analysis
II. Plot the stress-strain curves using MATLAB
III. Integrate the pristine stress-strain curve until the point of fracture to find the strain-energy density
IV. Determine the fracture energy by multiplying by the unstretched height of the sample

$$\Gamma = W(\varepsilon_c) \cdot H^*$$

**Relaxation and Creep Test:**
The crack length growth was measured through frame-by-frame video analysis and plotted over time.

*The formula for fracture energy, where $\Gamma$ = fracture energy ($J/m^2$), $\varepsilon_c$ = strain at rupture of pre-cut sample, H = unstretched sample height, $W(\varepsilon_c)$ = strain-energy density at strain $\varepsilon_c$

### 2. Experimentation

**Pure Shear Test:** comparison of a pre-cut and pristine sample at some loading rate to determine fracture energy
**Loading rates tested:** 10%/s, 1%/s, 0.1%/s, 0.01%/s

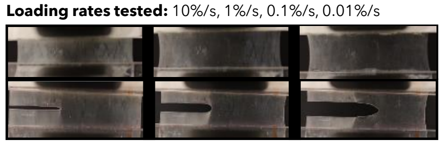


Figure 4. Pictures of a pristine and pre-cut LCE sample under a pure shear test

**Relaxation Test:** LCE sample held at a constant strain over time
**Strains tested:** 0.1, 0.15, 0.2, 0.27, 0.34

**Creep Test:** LCE sample held at a constant stress over time
**Stress tested:** 0.44 MPa

All tests conducted with a tensile testing machine (Instron 5944)

## Results

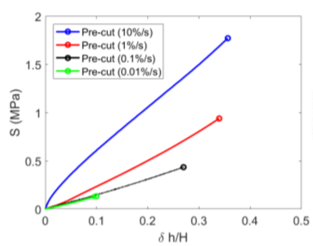



Figure 5. Stress-strain curves of pre-cut samples under the pure shear test

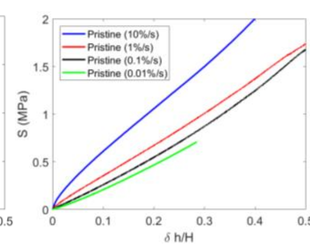



Figure 6. Stress-strain curves of pristine samples under the pure shear test

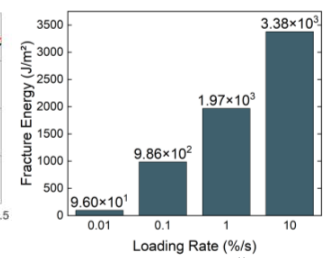



Figure 7. Fracture energy at different loading rates measured by pure shear test

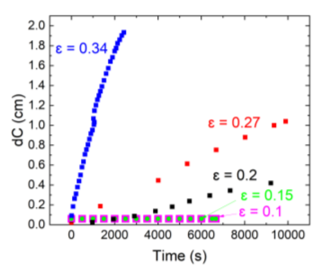



Figure 8. Crack propagation over time at different strains under the relaxation test

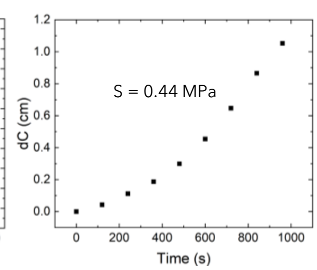



Figure 9. Crack propagation over time at 0.44 MPa under the creep test. 0.44 MPa corresponds to the critical stress value of the 0.1%/s pure shear test


Figure 10. Different crack tip shapes observed at fast and slow loading rates

## Conclusions

- Our experiments show that the fracture energy is directly correlated with the loading rate, which agrees with previous literature
- Comparison of the fracture propagation in the relaxation tests indicate that the threshold for delayed fracture is at approximately 2 mm of displacement (strain of 0.2)
- In the future, we will study fatigue fracture behavior as well as different fracture modes in LCEs

## References

[1] Saed, M. O., Torbati, A. H., Nair, D. P., & Yakacki, C. M. (2016). Synthesis of programmable main-chain liquid-crystalline elastomers using a two-stage thiol-acrylate reaction. *Journal of Visualized Experiments*, (107). https://doi.org/10.3791/53546-v

## Acknowledgements

I would like to acknowledge and thank Professor Jin and my daily lab supervisor Chen Wei for providing me with this opportunity. I would also like to acknowledge the SURP facilitators for organizing this program. This work is supported by the National Science Foundation (NSF) through Grant No. CMMI-1925790 and an NSF Research Experiences for Undergraduates (REU) supplement Grant No. CMMI-2327886.


## FACULTY ADVISOR

**Ximin He**

## DAILY LAB SUPERVISOR

**Zixiao Liu**

## DEPARTMENT

**Materials Science and Engineering**

## RYAN RUSCH

Materials Engineering
2nd Year, UCLA


# A Photothermal Oscillatory Triboelectric Nanogenerator

## ABSTRACT


Soft robotics is an innovative field at the intersection of engineering and biology, focusing on the development of robots composed of flexible materials. Unlike traditional rigid robots, soft robots mimic the adaptability and versatility of nat ural organisms, enabling them to interact uniquely in dynamic environments. Providing electrical energy to these robots typically demanded rigid elements like batteries or solar panels. Hence, a novel approach is needed to construct a fully self-powered soft robot. Recent literature has demonstrated the applica tions of a composite with a liquid crystal elastomer (LCE) and silicone rubber PDMS (polydimethylsiloxane). The LCE, doped with a light-absorbing agent will shrink in response to light, causing constant undulations when illuminated.

The conversion of these undulations to electrical energy can be accomplished using the triboelectric effect, where the contact-separation of PDMS and the nonstick material PTFE (polytetrafluoroethylene) build up electric charges on the surface of the materials to be measured as voltage. The experimental device was created by adhering with silicone glue the strips of material in the following order: LCE-PDMS-PTFE-LCE. Under illumination with a 780mW/cm2laser, undulations in the device arose, and an AC voltage of amplitude up to 50 mV and frequency up to 1 Hz was generated. These undulations mirrored those of stingrays, suggesting a new bio-inspired pathway for energy harvesting. Over all, we conclude that a multilayered design combining the LCE oscillator and the triboelectric properties of PTFE is a viable way of providing power to the autonomous soft robots of the future.

# A Photothermal Oscillatory Triboelectric Nanogenerator

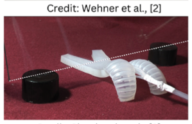
**Ryan Rusch, Zixiao Liu, Prof. Ximin He**
UCLA Materials Science & Engineering

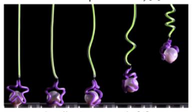
Samueli Research Scholars
UCLA Samueli School of Engineering

UCLA Samueli School of Engineering
SUMMER UNDERGRADUATE RESEARCH PROGRAM

## Introduction


Credit: Wehner et al., [2]


Credit: Shepherd et al., [3]


Credit: Jones et al., [4]

Soft robotics is an innovative field at the intersection of engineering and biology, focusing on the development of **robots composed of flexible materials**. Unlike traditional rigid robots, soft robots mimic the **adaptability and versatility** of natural organisms, enabling them to interact uniquely in **dynamic environments**. Providing electrical energy to these robots typically demanded rigid elements like batteries or solar panels. Hence, a novel approach is needed to **construct a fully self-powered soft robot**.

## Materials

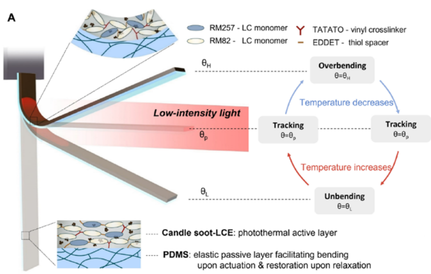
- **PDMS** (polydimethylsiloxane)
  - Elastic Silicone Rubber
  - Tribopositive
- **PTFE** (polytetrafluoroethylene)
  - Nonstick Fabric
  - Tribonegative
- **Candle Soot** doped **Liquid Crystal Elastomer**
  - Undulates under laser illumination
- **Silicone Glue**
- **Gold Electrode**

## Methods

1. **PDMS Synthesis** using Sylgard 184, 5:1 monomer/curing ratio
2. **Gold Deposition** onto PDMS and PTFE strips
3. **Adhesion** of LCE strips to PDMS and PTFE
4. **Adhesion** of PDMS to PTFE, leaving gap in the center
5. **Laser Illumination** parallel to apparatus, position and electrical data measured

## Background - Oscillator and TENG

### Liquid Crystal Elastomer (LCE) Oscillators

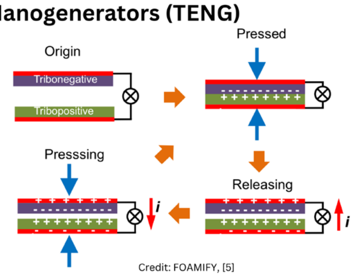


Credit: Shi et al, [1]

- Under illumination, the Liquid Crystal Elastomer (LCE) contracts, curling the device upwards.
- As the hinge is covered and cooled, it expands and the device curls downward
- Repeated curling is observed as oscillation

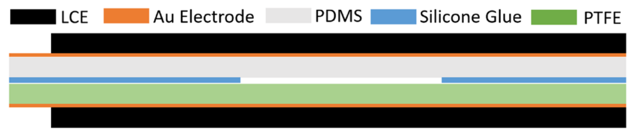
### Triboelectric Nanogenerators (TENG)

- Different materials contact and separate, gain charge
  - "Tribopolarity" determined by material work function
- Built up static charge induces opposing charge in electrodes, acts as capacitor
- Charge is collected as electric current



Credit: FOAMIFY, [5]

## The Apparatus




## Results

### Undulatory Behavior

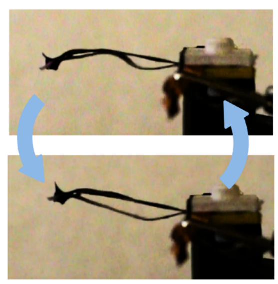


Figure 1: Two states of device, oscillated between during illumination

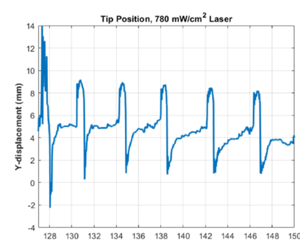


Figure 2: Position of the tip of the device under illumination from the 780 mW/cm^2 laser

- **"Manual Startup Oscillations"** required to convert "tracking" behavior into undulations
- **0.3 Hz Undulatory "snapping"** behavior observed in PTFE (bottom) composite
- **~50 mV spike** (open circuit voltage) observed alongside each "snap"
- Failure behavior:
  - **Adhesive Breakdown**, delamination of PTFE and PDMS
  - **Combustion of LCE** due to overheating

### Electrical Behavior

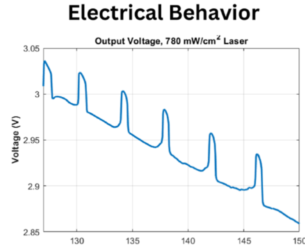


Figure 3: Open circuit voltage of the device under illumination from the 780 mW/cm^2 laser

## Conclusions

- A **multilayered design** of LCE-PDMS-PTFE-LCE is a potential design for **providing electrical power in soft robotics**
  - The LCE Oscillator design can be extended to include non-PDMS materials
  - PTFE-PDMS triboelectric pair is sufficient to extract electricity from motion
- Greater research adjusting parameters in fabrication (adhesive amounts, layer thicknesses, etc) is necessary to increase voltage, current, stability

## References

[1]: Shi, P., Zhao, Y., Liu, Z., He, X. Liquid crystal elastomer composite-based photo-oscillator for microrobots. *Journal of Composite Materials*. 2023;57(4):633-643.
[2]: Wehner, M., Truby, R., Fitzgerald, D. et al. An integrated design and fabrication strategy for entirely soft, autonomous robots. *Nature* 536, 451–455 (2016).
[3]: Shepherd, R.F., F. Ilievski, W. Choi, et al. Multigait soft robot. *Proceedings of the National Academy of Sciences of the United States of America*, 108(51), 20400-20403 (2011).
[4]:Jones, T.J., Jambon-Puillet, E., Marthelot, J. et al. Bubble casting soft robotics. *Nature* 599, 229–233 (2021).
[5]: Wicklein, B. "The TENG Working Principle." *FOAMIFY*.

## Acknowledgements

This project was funded by the Samueli Research Scholars program, directed by Associate Dean Veronica Santos. Research was completed in the lab of **Professor Ximin He** of UCLA's Materials Science & Engineering Department, under the direction of Daily Lab Supervisor **Zixiao Liu**.

## FACULTY ADVISOR
**Achuta Kadambi**

## DAILY LAB SUPERVISOR
**Achuta Kadambi, Sasha Vilesov and Pradyumna Chari**

## DEPARTMENT
**Electrical and Computer Engineering**


# SELIM EMIR CAN

Electrical and Computer Engineering
3rd Year, UCLA


# Enabling Diverse Eye Anatomy Tracking


## ABSTRACT

Eye tracking is a sensor technology that uses one or more cameras, infrared light sources, and computing capabilities to measure the movement of a person's eyes for applications such as: gaze estimation, disease diagnosis (ADHD,OCD, ASD, Parkinsons), eye health research, and surgical robotics. Particularly for gaze estimation, the performance of eye tracking critically depends on how well the test data distribution is covered by the calibration data. In practice recording calibration data that covers the full distribution is challenging and has been addressed by rendering photo-realistic images of the eye through the use of eye models generated using computer graphics techniques. However, the use of simplified eye models is one major limitation of previous work.

We hypothesize that eye tracking is a function of eye anatomy and that using a more anatomically correct and more anatomically diverse eye model result in improved eye tracking performance. An eye model that can be varied in eyeball size, corneal shape, corneal thickness, and iris size was created in Blender 3.6. Combined with a dense 3D head scan of 10 subjects, images of this eye model are rendered from different perspectives under different lighting conditions. We successfully modeled the refractive caustics of the cornea, created geometrically accurate sclera and cornea, and automated scene setup and eye model creation using scripting.We trained a learning-based gaze estimator on the rendered images of our model and evaluated the change in performance. In the future, we would like to iterate on our current eye model design to make it even more photo-realistic.

UCLA Samueli School of Engineering

SUMMER UNDERGRADUATE RESEARCH PROGRAM

# Enabling Diverse Eye Anatomy Tracking


Selim Emir Can, Alexander Vilesov, Pradyumna Chari | Achuta Kadambi

Department of Electrical and Computer Engineering

The Kadambi Lab / Visual Machines Group


## Background & Introduction

**Eye tracking** is a sensor technology that can detect a person's presence and follow what they are looking at in real-time.

- Need to capture the underlying distribution of data for reliable results
- Performance depends on how well the test data distribution is covered by the training set
- Data collection is expensive!

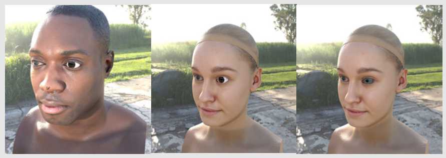


Fig 2. Integration of our eye model with remeshed dense 3D head scan of subjects showing diverse human physiology (left and middle image). The image on the right shows previous model which is not anatomically accurate (Only the left eye was replaced with eye models)

**Solution**: Render images from synthetic eye model!

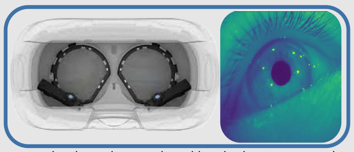


Fig 1. VR headset with eye-tracker add-on that has 1 camera and 2 infrared light sources per eye. The image on the left is typical data captured by eye tracker

### Applications

- Assistive technologies for disabled individuals
- Neurological Disorder Diagnosis
- Surgical Robotics
- Medical Education

### What can be improved?

Previous work doesn't use an anatomically accurate eye model creating a potential domain gap between real and synthetic data

**Hypothesis**: eye tracking is a function of eye anatomy and using a more anatomically correct/diverse eye model result in improved eye tracking performance.

## Key Terms

- **Gaze estimation** is the process of identifying the line of sight for each eye of a human user at a single instant
- **Rendering** is the process of generating a photorealistic or non-photorealistic image from a 2D or 3D model by means of a computer program.
- **Iris**: The colored membrane around your **pupil**.
- **Cornea**: The clear front outer layer of your eye. It covers the iris.
- **Sclera**: The outer coat of the eyeball that forms the whites of your eyes.

## Methods and Design

The eye model was created in Blender 3.6 (computer graphics software):

- model the cornea as a biconic surface for anatomical accuracy [2]
- Iris curvature depends on eye health, Iris can be flat or curved in our model [3]
- utilize dense 3D head scan of 10 subjects for diversity
- use physiological parameters reported in previous research
- models head-slippage by generating random jittering

Fig 3. Blender Logo

### Biconic Surface Equation

$$z(x,y) = \frac{\frac{x^2}{R_x} + \frac{y^2}{R_y}}{1+\sqrt{1-\frac{(1+Q_y)y^2}{R_y^2}-\frac{(1+Q_x)x^2}{R_x^2}}}$$

for each meridian and rotation angle ϕ:

- Rx, Ry are principal radii of curvature
- Qx, Qy are asphericity values

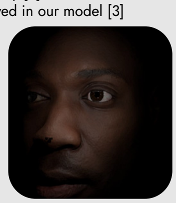


Fig 4. Image showing realistic lighting conditions for Virtual Reality data collection

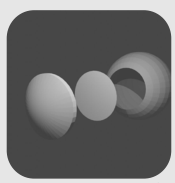


Fig 5. Deconstructed view of the mesh that makes up our eye model (cornea, iris, sclera from left to right)

corneal diameter: 11.79mm
anterior cornea measurements [2]
Rx: 7.63±0.29mm
Ry: 7.4±0.28mm
Qx: -0.46±0.14mm
Qy: -0.48±0.14mm

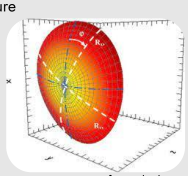


Fig 6. Biconic surface which is used to model the cornea

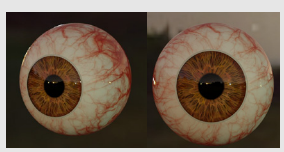


Fig 7. Rendered view of our standalone eye model (off-axis view on the left and on-axis view on the right)

## Results & Discussion

**Results**

- Modeled the refractive caustics of the cornea
- Created geometrically accurate sclera and cornea
- Automated scene setup eye model creation using python scripting

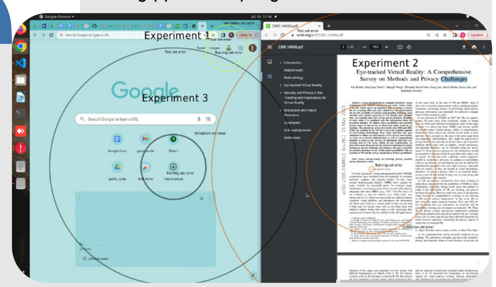



Fig 7. Visualizing gaze estimation error with real data only

- Trained a 6 layer convolutional neural network on the rendered images. Performance did not improve due to the following limitations:

**Limitations**

- Current camera configurations for rendering images does not precisely match the actual camera locations
- Textures need to be modeled with Infrared textures
- Encountered internal problems with unit conversion of ground truth labels which are important for training
- Skin texture can be improved to look more realistic

## Conclusion

### Why is this important?

To ensure consistent and reliable results for a diverse range of individuals, it's vital to create equitable eye tracking technology which is underscored by its clinical applications including the diagnosis of conditions like Parkinson's and assistive technologies such as gaze-based text entry for people with limited motor function (i.e. cerebral palsy). One of the main goals of my research is to reduce bias related to diverse eye structures and human physiology.

### Next Steps

- Will start using infrared textures instead of textures in the visible spectrum
- Will determine the exact location of eye tracking cameras to improve the quality of rendered images
- Will ensure that ground truth labels are in the proper units

## References


[1] Joohwan Kim, Michael Stengel, Alexander Majercik, Shalini De Mello, David Dunn, Samuli Laine, Morgan McGuire, and David Luebke. 2019. Nvgaze: An anatomically-informed dataset for low-latency, near-eye gaze estimation. In Proceedings of the 2019 CHI conference on human factors in computing systems. 1–12.

[2] Edgar Janunts, Marc Kannengießer, Achim Langenbucher, Parametric fitting of corneal height data to a biconic surface, Zeitschrift für Medizinische Physik, Volume 25, Issue 1, 2015, Pages 25-35, ISSN 0939-3889, https://doi.org/10.1016/j.zemedi.2014.02.005.

[3] Navarro R. The Optical Design of the Human Eye: a Critical Review. J Optom. 2009;2(1):3–18. doi: 10.3921/joptom.2009.3. Epub 2010 Nov 4. PMCID: PMC3972707.


## Acknowledgements


I would like to thank Prof. Achuta Kadambi, my PhD supervisors Sasha Vilesov and Pradyumna Chari as well as the SURP team and Samueli Research Scholars for funding my research.


## SETH FERRELL

Electrical and Computer Engineering
1st Year, UCLA



**FACULTY ADVISOR**

**Sudhakar Pamarti**

**DAILY LAB SUPERVISOR**

**Haris Suhail and Vinod Jacob**

**DEPARTMENT**

**Electrical and Computer Engineering**


# Testing and Analysis of a 0.2 to 5 GHz 4×4 MIMO Receiver Front-End RFIC Supporting Intra-Band Carrier Aggregation

## ABSTRACT


Mobile cellular phones, among many modern devices, require the use of multiple-in multiple-out (MIMO) receiver technology in order to communicate globally. These device's integrated receivers utilize off-chip Surface Acoustic Wave (SAW) filter technology to eliminate undesired signals. However, these filters take up expensive board real estate. Despite their high selectivity, SAW filters have fixed center frequencies and are unable to exceed frequencies above 3 GHz. By leveraging the filtering by aliasing technique, the SAW filter can be replaced by a highly linear adjustable bandpass filter that yields comparable outcomes in filtering interfering out-of-band signals (blockers) over 0.2 to 5 GHz enabling high dynamic range reception. This SAW-less front-end technology can aid in the future miniaturization of receiver design.

A radio frequency integrated circuit (RFIC) containing this new filter necessitates a robust testing circuit capable of emulating four simultaneous RF signals and blockers targeting the receiver inputs. A PCB was engineered in Altium Designer to drive and interface with the RFIC. The completed design included four 50 Ohm RF inputs, a differential local oscillator input, and a 10 GHz digital clock input. It was simulated within Keysight ADS in order to validate the driver's transmission lines' proficiency in minimizing reflections across a broad band, thus ensuring good signal integrity. Future work requires the design of an RF attenuator-based PCB capable of beamforming akin to that of a MIMO transmitter and the development of a python control program capable of testing the RFIC.

# Testing and Analysis of a 0.2 to 5 GHz 4×4 MIMO Receiver Front-End RFIC Supporting Intra-Band Carrier Aggregation

Seth Ferrell, Vinod Kurian Jacob, Haris Suhail, Prof. Sudhakar Pamarti
Department of Electrical and Computer Engineering

UCLA Samueli School of Engineering
SUMMER UNDERGRADUATE RESEARCH PROGRAM

## Background and Importance

- Mobile phones today carry multiple-in multiple-out (MIMO) receivers in order to **communicate globally**
- MIMOs implement Surface Acoustic Wave (SAW) and Bulk Acoustic Wave (BAW) filter technology to **block undesired signals**
- SAW/BAW filters have fixed center frequencies, are unable to exceed 3 GHz, and take up lots of expensive space
- With the filtering by aliasing technique, our lab developed a radio frequency integrated circuit (RFIC) to **replace the SAW/BAW filter** with an adjustable, 0.2 to 5 GHz equivalent filter
- RFIC front-end allows for intra-band carrier aggregation using two separate LO inputs

## Goals

1. Design PCBs to capture the capabilities of the 4x4 MIMO receiver front-end
2. Using Python, communicating via JTAG through an Arduino, perform tests and analysis on the RFIC

## Materials and Design

Altium Designer. KEYSIGHT TECHNOLOGIES. Advanced Design System

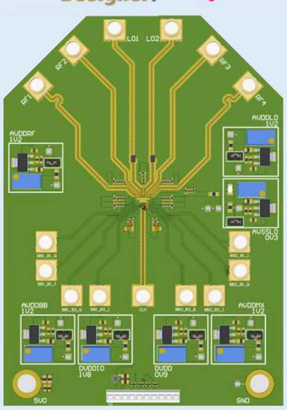


Figure 1. RFIC driver PCB.

Figure 1 shows a top-level view of the RFIC driver PCB. It contains:

- Four antenna inputs
- 2 single-ended LO inputs–on-board baluns convert them to balanced lines
- Up to 10 GHz clock input
- 7 voltage regulators
- 8 I/Q outputs for dual intra-band carrier aggregation

## Acknowledgements

I would like to thank members of the UCLA SPACE lab, with emphasis on V. Jacob and H. Suhail, for their help on this project. The RFIC being tested in this project is V. Jacob's work and I appreciate all the time he spent explaining its inner-workings. I would also like to thank SpectrumX for funding my research.

## Simulation Results

RF performance of the coplanar waveguides used on the PCB were characterized using ADS. Models designed in Altium were exported to ADS and EM-simulated using a 4 mil-thick Rogers 4350b substrate. Ideally, the return loss needs to be below −20 dB and the insertion loss must be near 0 dB.

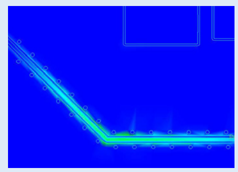


Figure 2. CPWG.

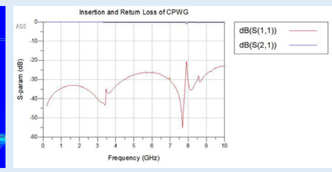


Figure 3. Simulation of CPWG.

Figure 2 shows a section of a coplanar waveguide to the input of the RFIC. Figure 3 shows a low $S_{11}$ (return loss) and low $S_{21}$ (insertion loss) at our frequencies of interest.

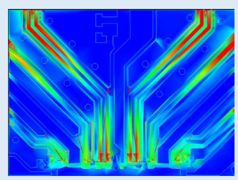


Figure 4. Coupled CPWGs.

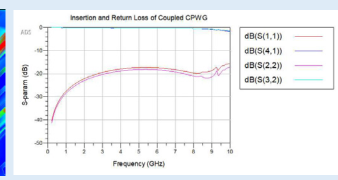


Figure 5. Simulation of coupled CPWGs.

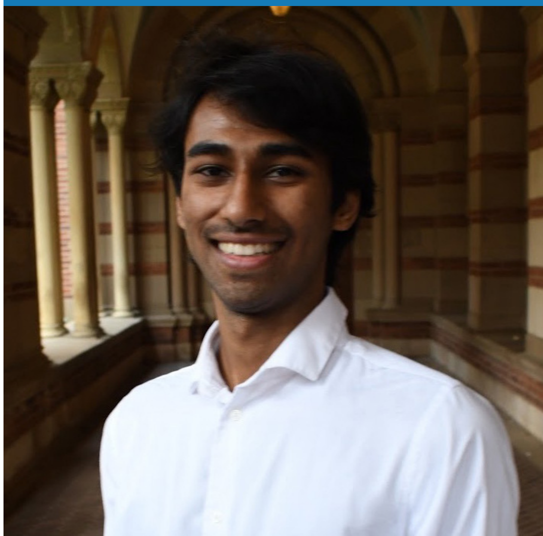
Figure 4 shows two coupled CPWG transmission lines (middle) used for the LO inputs. Figure 5 shows that both have a decent return loss over the frequency range of interest, but it diminishes above 3 GHz.

## Conclusion and Future Prospects

- Simulation results demonstrates that testing using this PCB is possible over the frequency range of interest
- Given enough time, future simulations will be done on the entire PCB to verify these results
- After manufacturing the driver, work will begin on the second PCB for generating the four RF inputs capable of beamforming with blocking signals like that of a real MIMO transmitter

## References

Golabighezelahmad, Sajad, et al. "A 0.7–5.7 GHz reconfigurable MIMO receiver architecture for analog spatial notch filtering using orthogonal beamforming." *IEEE Journal of Solid-State Circuits*, vol. 56, no. 5, 2021, pp. 1527–1540, https://doi.org/10.1109/jssc.2020.3028811.
Zhang, Linxiao, and Harish Krishnaswamy. "Arbitrary analog/RF spatial filtering for digital MIMO receiver arrays." *IEEE Journal of Solid-State Circuits*, vol. 52, no. 12, 2017, pp. 3392–3404, https://doi.org/10.1109/jssc.2017.2759118.



FACULTY ADVISOR
**Jonathan Kao**

DAILY LAB SUPERVISOR
**Jonathan Kao**

DEPARTMENT
**Electrical and Computer Engineering**


## SHREYAS KAASYAP

Electrical and Computer Engineering
1st Year, UCLA


# Using Surface Electromyography (sEMG) to Restore Autonomous Hand Control

## ABSTRACT


Brain computer interfaces (BCIs) are a rapidly growing field with the aim of creating machines that collect brain signals and translate them into desired actions. Applications of such work are widespread, from clinical work with paralyzed and amputated patients to creating smoother systems for virtual reality users. However, current state-of-the-art techniques require invasive placements of electrodes to record neural activity, which demands much and incurs significant risk for a patient.

My project focuses on using non-invasive surface electromyography (sEMG) signals that measure activity from motor neurons in the forearm. We use deep neural networks to decode these signals for applications such as real-time natural typing and cursor control, with goals of a functional prosthetic hand for amputated patients. The primary challenge is to train neural networks to correctly decode actions. This involves resolving meaningful neural signals in the presence of noise due to several factors, including interference from skin and interfering signals from other muscles in the hand. Furthermore, due to the physiological variability of humans, making such a network robust across patients provides a challenging task. We use convolutional neural networks to extract spatial-temporal information from the signal. For downstream tasks such as natural typing, we use a character-level language model in combination. We achieve 95% decoding accuracy in classifying 1 of 5 fingers, and can successfully play Pac-man and are working towards proficient typing. This work demonstrates that non-invasive sEMG can be used to play games and likely allow users to type naturally.

# Using Surface Electromyography (sEMG) to Restore Autonomous Hand Control

UCLA Samueli School of Engineering

SUMMER UNDERGRADUATE RESEARCH PROGRAM

Shreyas Kaasyap, John Zhou, Johannes Lee, Jonathan Kao
Department of Electrical and Computer Engineering, University of California, Los Angeles

## Introduction

- Brain Computer Interface (BCI): Machines that collect brain signals and decode them into desired actions

- Current state of the art techniques involve invasive placement of electrode arrays either on the brain or arm.

- Non-invasive surface electromyography (sEMG) involves taking electrical signals from the arm of a subject.

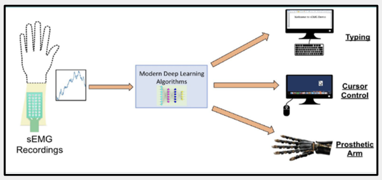



**Figure 1: Experimental Setup.**, High level view of project.

- Trade off is signal quality due to impedance and noise

## Methods and Materials

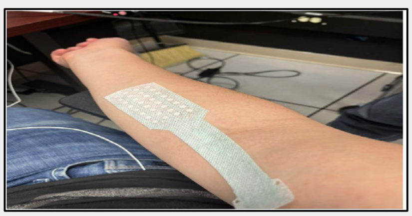


**Figure 2: Surface Electrodes**. 4x8 Electrode array placed on the skin. Electrodes measure electrical signal, corresponding placements of electrodes records different muscles.

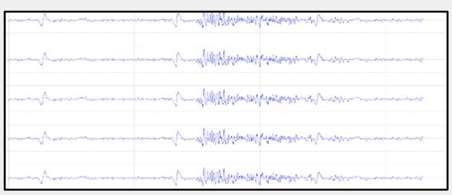


**Figure 3: Data Stream.** Filtered and plotted EMG activity stream. High amplitude activity represents activity.

- Electrical signals from forearm recorded from electrodes

- Noise from underlying muscle activity, activity is a combination of many motor unit action potentials

- Activity is heavily spatially-correlated and also heavily time-correlated

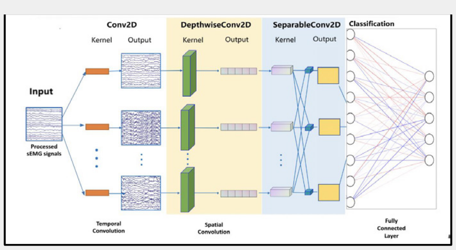



**Figure 4: Model Architecture**. We use a convolutional neural network to extract high-frequency features from the data. The input is a data stream that is viewed as an image along time. Convolutions are taken along spatial and temporal axes allowing feature extraction of spatial-temporal information in activity. This model architecture depicts fully connected layers that classify EMG signals into movement of one of the 5 fingers.

## Results

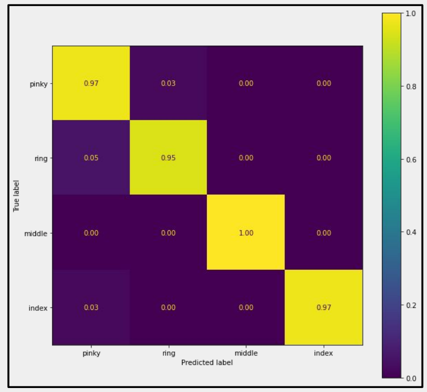



**Figure 5: Confusion Matrix of 4-finger task.** This confusion matrix shows the percentage of correctly predicted labels for each finger. The overall accuracy was 97.4%. This was taken from a 4-finger task used to train a model for the game Pac-man using the 'a', 's', 'd', and 'f' keys.

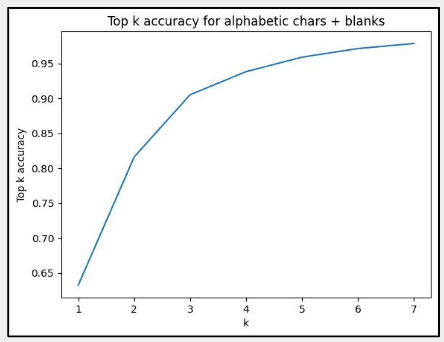



**Figure 6: Plot of top-k character decoding accuracy.** Plot was taken from a character decoding model, with 28 output classes (a-z, ' ', and a blank character for no keypress). For each k, if the true label is in the top k predicted output classes, it is counted as correct. We see a 65% character level accuracy with a 90% accuracy in the top 3 predicted labels.

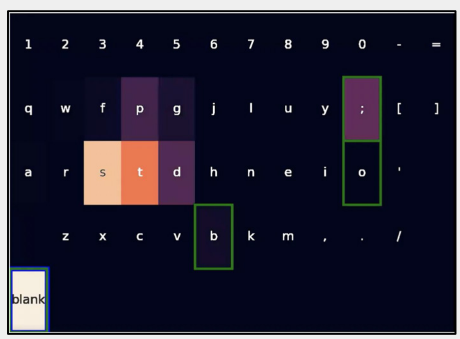


**Figure 7: Heatmap over keyboard.** Example of a system that is being used for closed-loop, real time solutions. Allows a user to visual the predicted outputs of the model at each timestep, and gives the user feedback over their presses.

## Discussion and Conclusion

We find high accuracy in finger decoding using sEMG and well above chance performance for character decoding. We have created real-time, closed-loop systems with low-latency decoding and have begun testing on paralyzed patients. We are currently incorporating LLMs into our decoders to work towards natural typing, and are excited to start testing on amputated patients. We hope to extend this project to gesture decoding and continuous control of the hand.

## Acknowledgments

I would like to thank Professor Jonathan Kao and the Neural Engineering and Computing Lab, as well as SURP and Fast Track for providing the resources for this project.


**FACULTY ADVISOR**
**Yang Zhang**

**DAILY LAB SUPERVISOR**
**Siyou Pei**

**DEPARTMENT**
**Electrical and Computer Engineering**

## SOURISH SASWADE

Electrical and Computer Engineering
1st Year, UCLA


# Using Machine Learning with Multimodal Sensor Fusion to Analyze Urban Road Conditions via e-Scooters

## ABSTRACT


Micro-mobility vehicles such as e-scooters and e-bikes are becoming increasingly prevalent, especially in urban areas where traditional forms of transportation like cars are inconvenient. Our system aims to (a) generate insightful conclusions on the real-time condition of roads and sidewalks by using electric scooters and (b) discover which modes of data (audio, visual, sensor) gathered from the scooter are most helpful to this goal. A mobile app continuously draws IMU (Inertial Measurement Unit), GPS, photo, and microphone data from a smartphone mounted on the electric scooter. Throughout a user’s ride, data is uploaded to and accessible from a cloud storage platform. Images are classified into road type (asphalt, brick, sidewalk) and road quality (smooth, worn, ruined) through two separate convolutional neural networks (CNNs). A third CNN extracts road quality from the audio recordings. Acceleration (IMU) time-series data determines road quality via a supervised learning model.

Road classification based on images offered the greatest accuracy (85-90%), making image data the most informative regarding road conditions, followed by audio data (likely due to intangible factors such as wind and human voices), and lastly, acceleration (IMU) data. Users on the system’s website select two coordinates on a map and receive data on the various road types and qualities along the route between the coordinates, aggregated from all modes of data. Applications of this system include aiding city infrastructure planners in prioritizing which streets to renovate, producing more convenient routes for wheelchair-enabled people by considering road slope, and helping users maximize their comfort while traveling.

# Using machine learning with multimodal sensor fusion to analyze urban road conditions via e-scooters

UCLA Samueli School of Engineering
SUMMER UNDERGRADUATE RESEARCH PROGRAM

Sourish Saswade, Yang Zhang, Siyou Pei
Department of Electrical & Computer Engineering, University of California — Los Angeles

## Introduction & Objectives

Every year, the usage of micro-mobility vehicles (MMVs) such as e-scooters and e-bikes becomes more prevalent, especially in urban areas where traditional forms of transportation like cars are inconvenient. Previous related research largely focused on bicycles rather than e-scooters, prioritized measuring the distance between MMVs and surrounding vehicles with expensive hardware, and collected but did not apply data to any practical scenarios.

By drawing and analyzing data from the IMU (Inertial Measurement Unit), camera, microphone, and GPS of a smartphone mounted on to an electric scooter, insightful conclusions can be made on the condition of roads and sidewalks the scooter rides on. This analysis is done through image, audio, and pattern analysis via machine learning models.

## Materials & Methods

**(1)** Equip e-scooter (Bird) with LG Nexus 5X using phone mount

User

Android Studio (Java) App

**(2)** Every GPS update, take a new photo and begin new microphone recording (~1 Hz frequency). Record orientation/rotation values at 10 Hz.

GPS Coord.
Rotation (IMU)
Acceleration (IMU)
Camera
Audio (Mic)

**(3)** GPS update → upload photo, previous recording to Firebase (~1 Hz frequency). Upload IMU values (10 Hz)

Firebase Firestore (GPS, IMU)
Firebase Storage (Camera, Audio)

**(4)** Download data

**Figure 1:** Data Acquisition platforms

Android Studio
Firebase

Google Colab Notebook (allows access to GPUs)

**Figure 2:** Data Analysis Platforms

PyTorch
matplotlib
NumPy

Image data → TRANSFORM: 1. RESIZE (to 224x224) 2. CONVERT TO TENSOR 3. NORMALIZE → ResNet-50 CNN, ResNet-50 CNN

Vertical Acceleration time series data → TRANSFORM: 1. FAST FOURIER TRANSFORM 2. PAD EACH SAMPLE TO LENGTH OF LONGEST SAMPLE → Support Vector Machine

Audio recordings → TRANSFORM: 1. CONVERT TO MELSPECTROGRAM TENSOR 2. PAD EACH SAMPLE TO LENGTH OF LONGEST SAMPLE → ResNet-34 CNN

ROAD TYPE classification (asphalt, sidewalk, brick)

3 ROAD QUALITY classifications/ratings **averaged** (ruined: 0, worn: 5, smooth: 10)

- Display data on React-built WEBPAGE with Google-enabled interactive map
- Map includes all coordinates for which we have road data
- Road classifications shown for each inter-coordinate stretch of road that user specifies

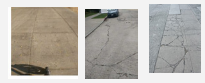
**Figure 3:** Data Display Platforms

React JS
Google Maps

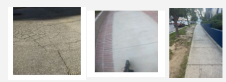
## Results

**IMAGE: 85-90% classification accuracy**

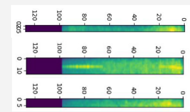
ROAD QUALITY (smooth→worn→ruined)



ROAD TYPE (asphalt →brick→sidewalk)

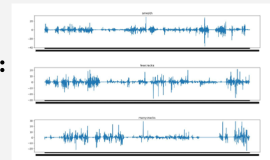


**AUDIO (mic): 68% classification accuracy**



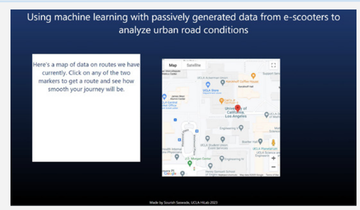
Example Spectrograms (smooth→worn→ruined, top to bottom)

**IMU (Vertical Acceleration): 53% classification accuracy**



All data for each classification, concatenated (smooth→worn→ruined, top to bottom)

**User Interface:** A marker on the map represents a coordinate — can click on two coordinates to receive data on the route between them




Road classification based on IMAGES offered the greatest accuracy (85-90%) for both type and quality, making image data the most informative regarding road conditions. Audio data is the next most informative, and acceleration time series data is the least.

## Conclusion & Future Additions

The image data offered a ~25% greater classification accuracy than the microphone recordings, and ~66% greater accuracy than the IMU time-series data. Audio recordings were likely affected by wind and human voices, as data was collected on real e-scooter rides. More precise calibration for the smartphone's acceleration measurements and a larger dataset could improve the IMU data's accuracy.

**USE CASES:**
(1) Aid city infrastructure planners in prioritizing which streets to renovate
(2) Produce more convenient routes for wheelchair-enabled people by considering road slope
(3) Help users maximize general comfort by recommending the route with the best conditions

**FUTURE ADDITIONS:**
(1) A more robust dashboard to display real-time/most recent changes in road conditions in a specified area, in addition to current map
(2) Utilization of LiDAR camera in more recent phones to precisely map cracks and defects in road — "An X-ray scan of the road"

## Acknowledgements

I would like to thank my PI Professor Yang Zhang, my daily supervisor Siyou Pei, the UCLA HiLab, and the UCLA Fast Track to Success Program for supporting this research. I would also like to thank the NSF (National Science Foundation) for funding this project.

A. M. Pérez-Zuriaga, D. Llopis-Castelló, V. Just-Martínez, A. S. Fonseca-Cabrera, C. Alonso-Troyano, and A. García, "Implementation of a Low-Cost Data Acquisition System on an E-Scooter for Micromobility Research," Sensors, vol. 22, no. 21, p. 8215, Oct. 2022, doi: 10.3390/s22218215.


## STEVE ZANG

Computer Science
2nd Year, UCLA

**FACULTY ADVISOR**
**Nader Sehatbaksh**

**DAILY LAB SUPERVISOR**
**Fatemeh Arkannezhad**

**DEPARTMENT**
**Electrical and Computer Engineering**


# Hybrid Attestation on Neorv32

## ABSTRACT


Remote attestation guarantees secure computing environments by authenticating the integrity of software running remotely. Among various solutions, methods based on hardware-software co-design (hybrid) are particularly popular due to their low overhead yet effective approaches. In this project, we implement a new hybrid attestation framework on NEORV32 RISC-V processors. NEORV32 is an open-source, customizable RISC-V processor. Users can personalize the core according to their specific needs by enabling/disabling relevant features.

In our threat model, we assume the adversary gains complete device control, modifying program and configuration data maliciously. Our solution is to utilize a software-hardware approach employing hash functions to identify these attacks. We present the implementation of the hybrid attestation on the NEORV32 RISC-V Processor, integrating the hash function hardware module, and ensuring its tampering resistance. The proposed hybrid attestation mechanism strengthens NEORV32-based device security and presents a valuable learning opportunity for developers and researchers interested in embedded system security. The code and data from this work are open source.

# Hybrid Attestation on NEORV32 RISC-V Processor

Steve Zang, Fatemeh Arkannezhad, and Nader Sehatbakhsh
Department of Electrical and Computer Engineering, University of California – Los Angeles

UCLA Samueli School of Engineering
SUMMER UNDERGRADUATE RESEARCH PROGRAM

## Introduction

**Neorv32**

- Open source, compatible, efficient RISC-V processor system
- Designed for embedded systems, IoT devices, and FPGA implementations
- Fully configurable and support hardware customization
- Operate in real time constraint environment, handle sensitive data

**Problem**

- No robust security mechanism exist for Neorv32
- Vulnerable to adversary attack

**Proposal**

- Implement Hybrid Attestation on NEORV32
- Executing only verified and legitimate software.

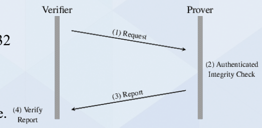



**Figure 1: General Remote Attestation Process**

## Background & Goals

**Background**

General Security Attestation:

- Prover
  Calculate Hash of its own memory and send it to verifier
- Verifier
  Trusted third party responsible for assessing measurements from prover

The prover generates the cryptographic evidence, and the verifier evaluates this evidence to verify the software's trustworthiness, providing an additional layer of security and assurance for the system.

**Research Goals**

Enable More Robust Security Mechanism by Implementing Hybrid Attestation on NEORV32

- Verify software integrity
- Prevent malicious code execution

## Experimental Setup

**Threat Model**

- Adversary gains complete device control, modifying program and configuration data maliciously
- Adversary doesn't modify the hardware of the device

**Overview**

- Implement the Hash Function in Hardware
- Customize the Custom Functional Subsystem (CFS)

**Original Proposal:**

- Software Attestation

**Problem:**

- Adversary may change hash function entirely
  → Not secured computation
- Use timing to detect genuinely
  → Uncertainty in threshold timing
- High CPU Cycle
  → User don't want to wait too long

**Current Proposal:**

- Hybrid Attestation

**Changes made:**

- Hardware customization
  → Implement hash function in hardware

**Expectation:**

- Overcome problems in Software Attestation

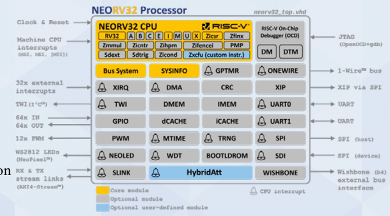


**Figure 3: Proposed Modification Based on NEORV32 Processor**

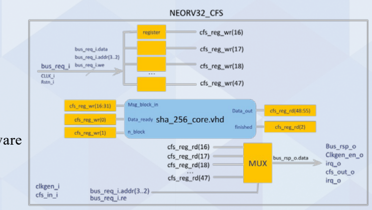


**Figure 4: Hardware Customization for Hash Function**

## Method

**Algorithm**

| Verifier V | | Prover P |
|---|---|---|
| Store S, (C, R) | | Store S, C |
| Generate IV | | |
| Save current time t | IV → | |
| $\sigma_0 \leftarrow IV$ | | $\sigma_0 \leftarrow IV$ |
| for i = 1,… , N do | | for i = 1,… , N do |
| Select Block $M_i$ | | Select Block $M_i$ |
| $\sigma_i \leftarrow$ HybridAtt $(S, \sigma_{i-1}, M_i)$ | | $\sigma_i \leftarrow$ HybridAtt $(S, \sigma_{i-1}, M_i)$ |
| end for | | end for |
| Save current time t′ | | $\sigma \leftarrow \sigma_N$ |
| if $(t - t' \leq \delta$ AND $\sigma = \sigma_N)$ Then | ← σ | |
| Accept | | |
| Else | | |
| Reject | | |

**Algorithm 1: Proposed Hybrid Attestation Technique.**

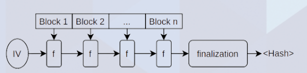



**Figure 2: Attestation Hashing Process**

**Attestation Process**

**Verifier (a trusted device):**

- (1) Send a challenge (IV) to the Prover to initiate the attestation process.
- (2) Verifier will also simultaneously continue the hashing process till it reaches to all blocks and keep track of its time to complete
- (4) Upon received answer from the Prover, compared with the one it had: if answer same —> compare time —> if time same —> remote device not compromised; else remote device compromised

**Prover (a remote device):**

- (2) Upon received the challenge, Prover will start the hashing process till it reaches all blocks
- (3) Record the answer and the timing it takes back to verifier

## Results & Overhead

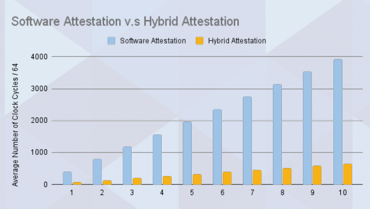



**Figure 5: Timing of Software Attestation and Hybrid Attestation.**

**Overhead**

**Hybrid Attestation:**

- Additional area cost for hash hardware module implementation
- Additional power consumed by hash hardware module

**Software Attestation:**

- Longer attestation cycle

**Notable Results**

**Attestation Timing:**

- Hybrid Attestation has an overall faster attestation clock cycles compared to Software Attestation

**Further Discussion:**

- Software Attestation:
  Hashing & Attestation Process complete in sequential order
- Hybrid Attestation:
  Hashing & Attestation Process complete in parallel order

**Memory Utilization**

| *text* | *data* | *bss* | *dec* | *hex* | *filename* |
|---|---|---|---|---|---|
| 8396 | 1072 | 116 | 9584 | 2570 | main.elf |

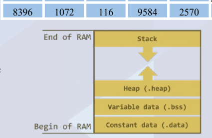



**Figure 6: RAM Layout**

## Conclusion and Future Work

- It's feasible to implement Hybrid Attestation on NEORV32 in real life since its overhead is small and it's a good tradeoff
- Utilize TRNG/PRNG to find set of address to be attested instead of going though the entire memory

## References

Surminski, S., Niesler, C., Brasser, F., Davi, L., & Sadeghi, A. (2021). *RealSWATT: Remote Software-based Attestation for Embedded Devices under Realtime Constraints.* https://doi.org/10.1145/3460120.3484788

Nolting, S., & All the Awesome Contributors. (2023). The NEORV32 RISC-V Processor [Computer software]. https://doi.org/10.5281/zenodo.5018888

## Acknowledgements

Special thanks to Professor Sehatbakhsh and Fatemeh Arkannezhad for their knowledge and support as well as the opportunity to work in the Secure Systems and Architectures Lab. We would also like to thank Will Herrera and the Summer Undergraduate Research Program staff for their guidance throughout the research process.

**FACULTY ADVISOR**
Suhas Diggavi

**DAILY LAB SUPERVISOR**
Dhaivat Joshi

**DEPARTMENT**
Electrical and Computer Engineering

**STEVEN LEWIS**

Electrical Engineering
2nd Year, UCLA

# De Novo Genome Assembly Using Current-level Modeling of Nanopore Reads

## ABSTRACT

Accurate genome assembly is impeded by the complexity of genomic repeat regions and the high error rates of long-read sequencing technologies. Nanopore sequencing, which produces long-reads, is particularly advantageous as it can sequence reads that are larger than repeat regions, giving us a more comprehensive view of the genome than its short-read counterparts. Short-read assemblies struggle to accurately resolve complex repetitive sequences leading to incomplete or erroneous reconstructions of genomic architecture. However, the use of nanopore data in assembly brings about its own challenges, specifically higher error rates and sequencing biases. Even state-of-the-art assembly algorithms like Flye face limitations in addressing these challenges.

Our research investigates the application of current-level modeling of nanopore sequencing in synergy with Flye. By leveraging the error biases in nanopore sequencing and applying quantization to the current-level signals, we hypothesize that we can significantly improve assembly accuracy, especially in these complex repeat regions. The proposed approach could allow Flye and similar assemblers to work more efficiently and assemble intricate repeat regions with greater precision. Utilizing nanopore long-reads of E. coli and D. melanogaster genomes, we benchmarked our Q3 assembly approach against the traditional nucleotide assemblies of raw nanopore reads data. The Q3 assemblies have shown intriguing patterns in assembly contiguity and accuracy as compared to the nucleotide assembly. As we continue to implement these methods, we anticipate improvements in assembly accuracy, which will be validated using multiple benchmarks from Flye and QUAST.

# De Novo Genome Assembly using Current-level Modeling of Nanopore Reads

Steven Lewis[1], Dhaivat Joshi[1], and Suhas Diggavi[1]
[1]Electrical and Computer Engineering Department

UCLA Samueli School of Engineering
SUMMER UNDERGRADUATE RESEARCH PROGRAM

## Objective

We aim to enhance the accuracy of genomic assembly, particularly in complex repeat regions, by integrating current-level modeling of nanopore sequencing and quantization with assembly algorithms like Flye, thus capitalizing on the error biases inherent in nanopore sequencing.

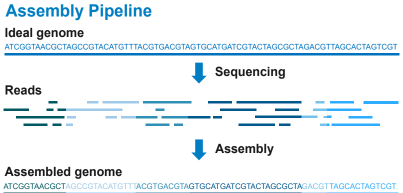
## Background

Complex repeat regions and high error rates in long-read sequencing technologies hinder accurate genome assembly.

### Motivation for Genome Assembly

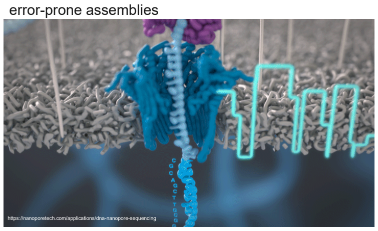
- Personalized healthcare and earlier disease/cancer diagnosis and prevention
- Low-cost, fast genetic heritage understanding
- Detection of structural variants and single nucleotides in genomic structure

### Assembly Pipeline




### Nanopore Sequencing

- Enzyme guides DNA strands through a nanopore, with resulting current changes base-called to identify underlying base pair sequences
- Generates long reads (~10,000 to ~100,000 bp) that cover larger repeat regions
- High error rates (5 to 10%) and sequencing biases lead to error-prone assemblies



*DNA strand being unwound and fed through a nanopore, producing current levels*

### Flye: Graph-based De Novo Assembly Algorithm

- Critical pipeline component that assembles de novo long-read sequences, especially repeat regions, with high accuracy
- Utilizes an assembly graph-based approach

## References

1. Joshi, D., Diggavi, S. et al. (2023). HQALIGN: Aligning Nanopore Reads for SV Detection Using Current Level Modeling.
2. Kolmogorov, M. et al. (2019). Assembly of long, error-prone reads using repeat graphs. Nat Biotechnol, 37, 540–546.
3. Mikheenko, A. et al. (2018). Versatile genome assembly evaluation with QUAST-LG. Bioinformatics, Volume 34, Issue 13, i142–i150.

## Methodology

**ACGTAACGTATTG**

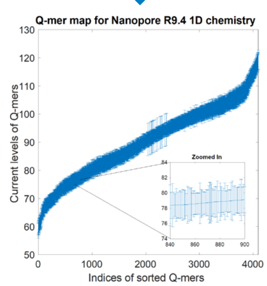
**Raw Nucleotide Sequencing Reads Data**

↓

**[ ACGTAA, CGTAA, GTAACG, TAACGT, AACGTA, ACGTAT, CGTATT, GTATTG]**

**List of 6-mers (Q-mers, Q=6)**

↓




*Q-mer map for MinION R9.4 Nanopore flowcell (Q = 6) showing significant overlap between adjacent Q-mers*

**Q-mer Map**

- Represents median current levels and error bounds for sequences of Q consecutive nucleotides (Q-mers) in nanopore sequencing
- **Captures inherent error biases in sequencing**
- **Highlights overlaps in current levels, key to understanding base-calling errors**

↓

**[ 82.13, 104.67, 81.75, 88.27, 97.85, 80.16, 104.75, 77.26 ]**

**List of Median Current Levels**

↓

**12011020**

**Q3 Quantized Sequence**

- Scheme maps ranges of current levels to 0, 1, 2 nanopore sequencing
- Thresholds are determined from entropy calculations
- **Has lower edit distance as compared to only assembling the raw base-called reads**

↓

**ACGTAACGTATTG**

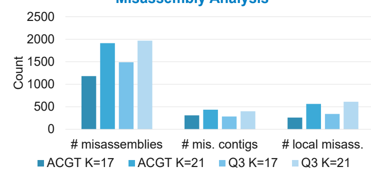
**Final Optimized Flye Assembly**

- Assembled with Flye using a memory-efficient k-mer counting algorithm used to count larger k-mers (k=21)
- **Larger k-mers are used to assemble complex repeat regions with greater accuracy**

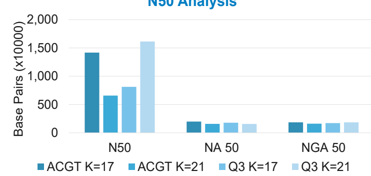
## Results

| | ACGT | | Q3 | |
|---|---|---|---|---|
| | k=17 | k=21 | k=17 | k=21 |
| Total Length | 139.535 M | 147.548 M | 142.428 M | 152.874 M |
| Genome % | 93.15% | 93.45% | 93.37% | 93.72% |
| # Contigs | 978 | 1133 | 734 | 1095 |
| Largest Contig | 24,386,828 | 24,033,285 | 24,417,534 | 24,223,356 |

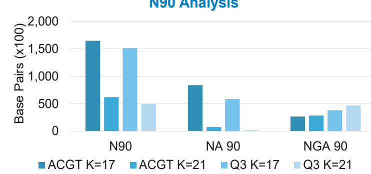
**Comparison of Assembly Quality Metrics: Misassembly Analysis** (A)




**Comparison of Assembly Quality Metrics: N50 Analysis** (B)




**Comparison of Assembly Quality Metrics: N90 Analysis** (C)




*Benchmarked using D. melanogaster dataset and QUAST v5.2.0*

## Discussion

- Q3 outshines ACGT in contiguity but has significant accuracy challenges
- There's an evident trade-off between assembly size and accuracy, notably in Q3 k=21
- Future strategies might optimize k-mer sizes or merge ACGT and Q3 methods for balance

### Overall Assembly Analysis

- All runs' largest contig is ~24 M bp; Q3 k=17 leads
- All runs cover ~93% genome; Q3 k=21 slightly ahead

### Misassembly Analysis (A)

- Q3 k=21 has most misassemblies, raising accuracy concerns
- Q3 k=17 has fewer, but more misassembled contigs, indicating quality issues
- Both methods show more misassemblies at k=21, suggesting k-mer size challenges

### N50 Analysis (B)

- ACGT k=17 and Q3 k=21 boast high N50 for contiguity
- Corrected values (NA50, NGA50) drop for ACGT k=17; Q3 k=21 excels in NGA50

### N90 Analysis (C)

- ACGT k=17's NA90 and NGA90 drop significantly
- Q3 k=17 maintains balance in all N90 metrics; Q3 k=21's low NA90 highlights its accuracy challenge

UCLA Samueli School of Engineering
SUMMER UNDERGRADUATE RESEARCH PROGRAM

Special thanks to Prof. Suhas Diggavi and Dhaivat Joshi from UCLA's Information Theory and Systems Laboratory for their valuable guidance. I would also like to thank my SURP mentors for supporting this project's organizational pieces.


## SUNDI WIN

Mechanical and Aerospace Engineering
2nd Year, UCLA

**FACULTY ADVISOR**
**Sungtaek Ju**

**DAILY LAB SUPERVISOR**
**Anthony Mannion**

**DEPARTMENT**
**Mechanical and Aerospace Engineering**


# Analysis of Lumped Capacitance Method for Mean Radiant Temperature


## ABSTRACT

Mean radiant temperature (MRT) is a crucial parameter for quantifying the effects of solar and surface radiation on human thermal comfort, especially in the context of climate change. The traditional methods of measuring MRT require either a globe thermometer, which has been proven to be inaccurate in outdoor spaces, or the inaccessible six-directional pyrgeometer and pyranometer setup. To address these limitations, this project investigates the viability of employing cost-effective sensors for MRT measurement. Specifically, it aims to extract a heat transfer coefficient utilizing the lumped capacitance approximation through the use of a reduced copper sphere. The sphere is exposed to solar radiation until it reaches a steady-state temperature, and is subsequently cooled by shielding. The time constant and consequent convective heat transfer coefficient is then determined by analyzing the cooling curves.

From initial results, we were able to derive a heat transfer coefficient of about 13-15 for W/m2K wind speeds of about 5-10 mph, respectively. Comparing these values to approximate heat transfer coefficients over a sphere of 33.5-39 W/m2K for similar air flow there is substantial error. While further investigation is still necessary to conclude the viability of this application of the lumped capacitance approximation, these results establish the groundwork for this method of MRT measurement. Future work includes data collection with a black sphere for maximum radiation absorption and also calculation of MRT values from the derived heat transfer coefficients to be compared with pyrgeometer measurements.

# Analysis of Lumped Capacitance Method for Mean Radiant Temperature

UCLA Samueli School of Engineering
SUMMER UNDERGRADUATE RESEARCH PROGRAM

Sundi Win, Anthony Mannion, Sungtaek Ju

*Department of Mechanical and Aerospace Engineering, University of California-Los Angeles*

## Introduction

**Mean radiant temperature (MRT)** : Uniform temperature of an imaginary enclosure in which radiant heat transfer from the human body equals radiant heat transfer in actual non-uniform enclosure[1]

- More accurate measurement of the way humans feel as human skin is more sensitive to radiant heat

**Lumped Capacitance Method:** Transient analysis method based on assumption that temperature gradient within solid is negligible

### Previous Methods

**Globe Thermometer**

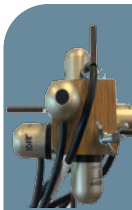
150 mm diameter

↓

Long equilibrium time (~20-30 min), not optimal for outdoor weather

1. Mitchell Instrument, *Globe thermometer*

**Net Radiometers**



Measure short and longwave radiation in 6 directions simultaneously

↓

Expensive and complex (integral radiation measurement, IRM)

2. Teitelbaum, *Pyrgeometers*

## Materials & Methods

- **Type K thermocouple** (TC) probes .020" diameter
- 3 cm hollow **copper sphere**

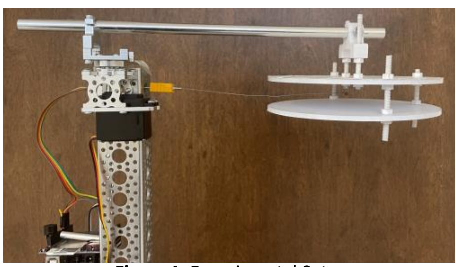


**Figure 1**: Experimental Setup

- 6 in 3D printed white solar **radiation shield** plates with **Tyvek** (~90% reflective fabric) cover
- **Arduino Uno** for data collection
- **MATLAB** for data processing and figures

### Methods

1. Allow copper sphere to heat up under sun
2. Radiation shield covers sphere to stop absorption and cool by convection
3. Repeat above process
4. Extract time constant ($\tau$) from cooling curve from exponential fit below

$$\frac{T - T_\infty}{T_i - T_\infty} = \exp\left(-\frac{\tau}{t}\right) \quad (1)$$

5. Calculate heat transfer coefficient from time constant

$$\tau = \frac{\rho V c}{h A_s} \quad (2)$$

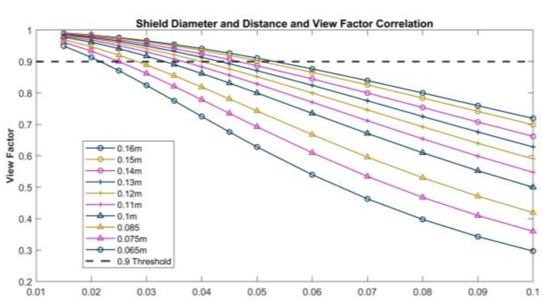



**Figure 2**: Plot of acceptable view factors for specified shield size and distance

## Results

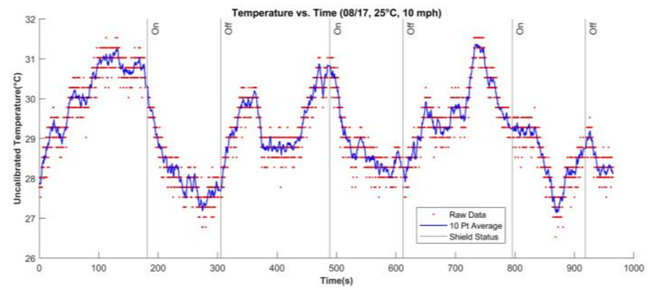
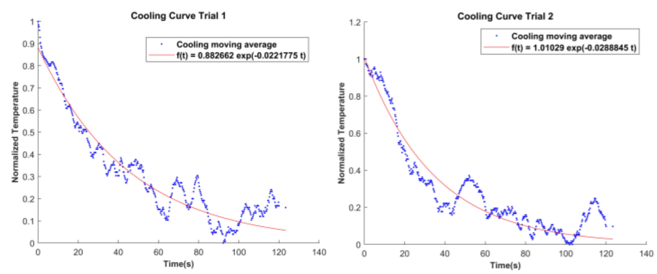



**Figure 3**: Raw data output from Arduino with 10 pt moving average




**Figures 4 (Left) & 5 (Right)** Exponential fitted curve (1) on moving average for respective cropped cooling trials

**Table 1**: Heat transfer coefficients at respective wind and temperature

| Wind Speed (m/s) | Temperature (°C) | Experimental Heat Transfer Coefficient ($W/m^2K$) | Calculated Heat Transfer Coefficient ($W/m^2K$) |
|---|---|---|---|
| 5 | 25.56 | 11.48~13.16 | 33.54 |
| 9 | 25 | 16.63~30.39 | 46.05 |
| 10 | 25.9 | 34.96~49.67 | 48.77 |

## Conclusion

As seen in Table 1, the experimental heat transfer coefficients range between 1.8-65.8% error of the calculated coefficients. While further investigation is still needed, these results demonstrate prospective viability in the lumped capacitance method for measuring MRT.

### Future Work

- Further data collection with anemometer and more extreme weather conditions
- Calculation of MRT compared to measurement from net radiometer using IRM method for accuracy
- Validate MRT with pyrgeometer and IRM method

## References

1. ISO, I. 7726, ergonomics of the thermal environment, instruments for measuring physical quantities. Geneva: Int. Standard Organ. (1998).

## Acknowledgements

I would like to thank the SURP program for providing the opportunity to conduct research this summer, including Will Herrera, Sam Negrete, and Amir Ahmadian. Thank you to Prof. Ju and Anthony for the continued support and guidance. Special thanks to Benjamin Tan for the technical support and expertise.


## TONG XIE

Applied Mathematics
2nd Year, UCLA

**FACULTY ADVISOR**
**Cho-Jui Hsieh**

**DAILY LAB SUPERVISOR**
**Andrew Bai**

**DEPARTMENT**
**Computer Science**


# Exploring Memorization in Diffusion Models: Tracing Influential Training Data

## ABSTRACT


Deep generative models are powerful tools that use neural networks to learn underlying data distributions and generate high-quality samples. However, memorization where the model fails to locally generalize to certain regions of the input space remains to be a challenge, leading to privacy and legal concerns. In particular, state-of-the-art diffusion models are shown to be much less private and more vulnerable to attacks. This research aims to study memorization in diffusion models by tracing model behavior to training data, and examining the most influential samples in shaping the performance and output of the diffusion process. We present an extension of the TracIn framework to diffusion models, enabling an efficient first-order approach to approximate the influence of training data in an unsupervised generative setting.

We demonstrate that samples with the highest self-influence manifest as high-contrast or visually distinct outliers, and samples with the highest test-influence are visually similar to the associated test sample. Furthermore, we leverage the extended TracIn framework to better understand the intricate learning dynamics of training samples across varying stages of the diffusion process. Through our study, we contribute to the interpretability of generative models and provide insights into the role of training data in model behavior, generalization, and memorization, paving the way for more transparent training techniques.

# Diffusion Models: Tracing Influential Training Samples

**Tong Xie[1], Andrew Bai[2], Cho-Jui Hsieh[2]**

[1] Department of Mathematics, [2] Department of Computer Science, University of California - Los Angeles

UCLA Samueli School of Engineering

SUMMER UNDERGRADUATE RESEARCH PROGRAM

## Introduction

Deep generative models are powerful tools that learn underlying data distributions and generate new high-quality samples.

In particular, the state-of-the-art diffusion models show exceptional capabilities in generating coherent images from pure noise, outperforming GAN and VAE.

However, these deep neural networks are often less understood and referred to as "black-box" methods, due to their complex interactions that drive pattern learning from training data.

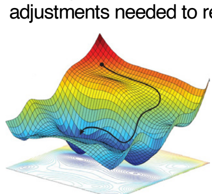
## Objective

This research traces model interpretability to training data by quantifying the influence of each sample on model performance. We accomplish this goal by extending the TracIn framework to diffusion models, offering an efficient instance-based interpretation in the unsupervised generative setting.

## Mathematical Intuitions

In training, model minimizes "**loss function**" that measures performance error

**Gradient** of loss function points to direction of steepest increase, shows the parameter adjustments needed to reduce loss



$z$ = training sample
$z'$ = test sample
$w$ = parameters
$l$ = loss function

**TracIn** traces how the interested training sample affects the loss on a test sample:

01 First-order approximation of how changes in parameters affect model's loss

02 Aggregate results over every saved checkpoints for each training-test pair

SGD update

$$w_{t+1} - w_t = -\eta_t \nabla l(w_t, z_t, T_t)$$

$$l(w_{t+1}, z') = l(w_t, z') + \nabla l(w_t, z') \cdot (w_{t+1} - w_t) + O(\|w_{t+1} - w_t\|^2)$$

$$\texttt{TracInCP}(z, z') = \sum_{i=1}^{k} \eta_i \nabla \ell(w_{t_i}, z) \cdot \nabla \ell(w_{t_i}, z')$$

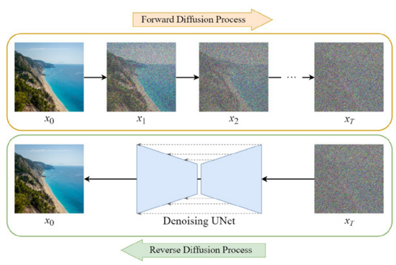
## Materials & Methods

### Materials

1. PyTorch: library for neural network based deep learning models
2. Oxford 102 Flower: dataset containing 102 flower categories, 40-258 images each
3. Model trained for **600 epochs** to transform pure noise into coherent image
4. Total of **15 checkpoints** used for TracIn

### Diffusion Model Architecture

1. UNet designed for image generation task (residual blocks, attention layers, sinusoidal time embeddings, down/upsampling)
2. 50 DDIM steps used for efficient sampling




*Examples of generated images trained for 200 epochs*

## Results

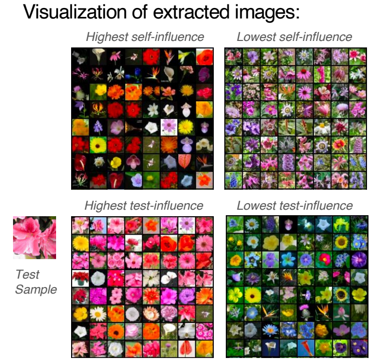
Gradient for test sample is averaged over 20 timesteps to approximate expectation of influence; we observe that the choice of timesteps significantly affect influence:

| Region | Top 1 | Top 3 | Top 10 |
|---|---|---|---|
| Early (1-200) | 33.6% | 58.6% | 82.8% |
| Middle (400-600) | 62.5% | 80.5% | 95.3% |
| Late (800-1000) | 7.8% | 21.1% | 41.4% |

Results for sanity check on the frequency that a training sample is most influential, among top 3, and among top 10 (using a subsample of size 128)

1. Using intermediate timesteps (400-600) gives the best performance extracting most influential samples
2. Later timesteps (800-100) gives less accurate measure of influence

Visualization of extracted images:




## Discussion

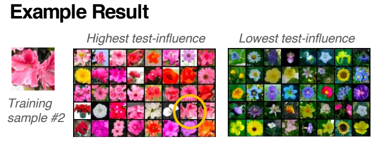
*Does this implementation of TracIn find the most influential training samples?*

**Sanity check**: training samples should have large influences over themselves

1 → 2 → 3

1. Compute test influence over a training sample *x*
2. Check if *x* is most influential among all samples over itself
3. *x* high ranking = successfully extracts influential samples

**Example Result**




Ranked 31 / 8176 samples, shows relatively high influence over itself

## Conclusions

**Timesteps & Learning Dynamics**

At intermediate diffusion stages, the model strikes a balance between detecting useful signal in noise and stabilizing behavior, thus using 20 timesteps from range 400-600 improves accuracy of calculated influence.

Earlier timesteps (1-200) are stages that contain more information about the coherent image, thus also provide satisfactory results.

**Extracted Samples**

The high / low influence training samples calculated by TracIn are visually intuitive:

| | Self-Influence | Test-Influence |
|---|---|---|
| High | High-contrast, visually distinct outliers | Visually similar to the associated test sample |
| Low | Low-contrast, ordinary colors and shapes | Visually dissimilar color, shape, or background |

## Future Works

1. Closely examine the effect of timesteps and correlation to learning dynamics at varying stages of diffusion process
2. Compare results to other measures of influence, such as influence function and representer method
3. Extend application of TracIn to outlier detection and memorization mitigation

## Acknowledgements

This work was supported by the National Science Foundation through UCLA Summer Undergraduate Research Program. I would also like to thank Professor Cho-Jui Hsieh and lab advisor Andrew Bai for providing guidance and mentorship throughout the project. A special thanks to Will Herrera and all SURP staff for organizing the program.

## Main References

[1] Jonathan Ho, et al. "Denoising Diffusion Probabilistic Models."
[2] Jiaming Song, et al. "Denoising Diffusion Implicit Models."
[3] Garima Pruthi, et al. "Estimating Training Data Influence by Tracing Gradient Descent."
[4] Zhifeng Kong, et al. "Understanding Instance-based Interpretability of Variational Auto-Encoders."

## FACULTY ADVISOR
**Yang Zhang**

## DAILY LAB SUPERVISOR
**Siyou Pei**

## DEPARTMENT
**Electrical and Computer Engineering**


## WILLIAM HUANG

Electrical and Computer Engineering
4th Year, UCLA


# Users in Wheelchairs (UIW) – A Human Centered RGB Dataset of Wheelchair Users


## ABSTRACT

Kinematic analysis of the human body has become an integral tool in multiple fields spanning from sports performance, rehabilitation, animations, and physical modeling. While traditional marker-based solutions are commonly used in kinematic analysis, it is not suitable for "in field" tasks. This problem is further exacerbated in users with disabilities, who may face difficulties in data collection settings. Markerless video-based motion capture systems using deep learning to estimate the poses of users through camera images, offering a more accessible alternative. These systems rely on large amounts of image data of humans in different day to day actions. However, most commonly available markerless motion capture systems including MediaPipe, OpenPose, BlazePose, and Detectron2 are purely trained off able-bodied datasets like COCO and consequently perform worse on users in wheelchairs.

To address this disparity, we present Users in Wheelchairs (UIW), a new dataset focused on wheelchair users. This is achieved by collecting data on common and extreme wheelchair related activities which are then placed in the context of real-world settings through manual video capture. Over 2.5k images were collected from 20 hours of video of wheelchair users in 21 different action classes from novel data collection software designed to capture and label video frames. Each instance's keypoints were then labeled through crowd worker involvement. We provide statistical analysis on our dataset in comparison to the COCO dataset and found a 32.50 AP improvement in bounding boxes on users in wheelchairs and a 13.34 AP improvement in keypoint precision.

UCLA Samueli School of Engineering
SUMMER UNDERGRADUATE RESEARCH PROGRAM

# Users in Wheelchairs (UIW) — A Human Centered RGB Dataset of Wheelchair Users

William Huang, Sam Ghahremani, Siyou Pei, Yang Zhang

## Introduction & Background

**Able-Bodied Users** - Users with no physical disabilities. For the context of this research, focused on users who are not in a wheelchair.

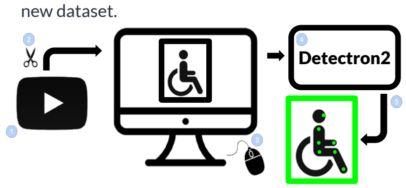
**Markerless Motion Capture (MMC)** - Motion capture systems, that rely on deep learning models to identify joint positions from camera images.

**Kinematic analysis** from **MMC** has played a key role in fields including rehabilitation, sports performance and robotics in large part due to its accessibility and cost. **MMC** is severely limited by the quality of the data used in training. Common algorithms (BlazePose, OpenPose, Detectron2) are trained off large **able-bodied** datasets like COCO. The lack of wheelchair users in these data sets leads to a disparity in performance between able-bodied users and users in wheelchairs.

**Figure 1:** Example output of OpenPose, a common MMC algorithm

## Objective

- Develop and collect a new dataset of RGB images of users in wheelchairs.
- Dataset is large and diverse enough to capture idiosyncrasies of movement in wheelchairs and the unique settings users in wheelchairs may be found.
- Validate new dataset to identify performance changes when testing on users in wheelchairs with new dataset.




1. Select a set of wheelchair videos to annotate from publicly available sources.
2. Collect frames with unique poses for annotation.
3. Annotate data through crowd worker involvement.
4. Boost existing Detectron2 object and pose detection models with data.
5. Validate performance changes of boosted model.

## Defining Action Groups

- In order to maintain consistency, define a set of actions and motions to collect videos of.
- Identified set **of common tasks both able-bodied and wheelchair users perform**.
- Selected most popular **wheelchair sports** for unique poses that may not be captured in everyday motions.

**Daily Life Examples**
- Stretches
- Daily Chores
- Wheelchair Skills

**Sports Examples**
- Basketball
- Rugby
- Dance

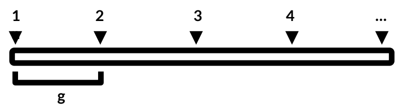
## Collecting Videos

- Selected **83 videos** in corresponding action groups.
- Collect from existing datasets and YouTube.
- Individual videos are labeled as the corresponding action group for crowd worker efficiency.

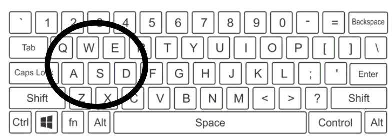
**Example Video Topics**
- Top 3 Stretches for Someone in a Wheelchair
- Shooting a Basketball From a Wheelchair
- Wheelchair Dance

## Frame Selector Application

- All videos must be analyzed manually before sending to crowd workers to select individual frames.
- Annotator is looking for motions and poses that sufficient differ from previously collected frames.
- Videos are collected through a custom multithreaded video scraper.
- Naive algorithm splits video into frames for annotator to analyze.




- For each video, select $n=500$ frames with a minimum of $g=60$ frames between each frame.
- If there are not $g*n$ frames in the video, select as many frames with g frames between each.




Each frame is displayed before the annotator has a set of options. Key bindings are designed for efficiency.
- D - Go to next frame
- W/S - Mark current frame as unique enough to save or not save. Defaulted to false.
- Q/E - Mark current frame as a frame with multiple people. Defaulted to false.

## Collected Frames

- Collected **2491 frames** from all videos.
- Each frame is labeled with the corresponding action of the video and whether knot the frame contains multiple people for crowd worker reference.

| Action Type | Number of Frames | % of Dataset |
|---|---|---|
| Talking | 458 | 18.38% |
| Daily Routine | 285 | 11.44% |
| Basketball | 231 | 9.27% |
| Dance | 225 | 9.03% |
| Wheelchair Skills | 171 | 6.86% |
| Moving | 153 | 6.14% |
| Tennis | 130 | 5.22% |
| Extreme Sports | 119 | 4.78% |
| Household Chores | 70 | 2.81% |
| Other | 295 | 11.84% |
| Total | 2491 | |

**Figure 2:** Table of distribution of actions in dataset

## Crowd Worker Validation

- Crowd workers are recruited to annotate keypoints and bounding boxes around users in wheelchairs.
- Some images are misannotated and must be manually reviewed and corrected.

**Figure 3:** Example of incorrect annotation. Knee is duplicated.

**Figure 4:** Example of correct annotation.

*Images from MobilityAIDS [3], a publicly available disability dataset. Labels by our crowd workers*

## Results & Discussion

Detectron2 was selected as the MMC model for testing. All models were boosted from the base COCO R50 model trained on the COCO Keypoint dataset. UIW was separated into a 1.5k training set and trained for 30 epochs with default hyperparameters and a learning rate of .00025 tested on a 1k testing set. Base COCO R50 was used as the control.

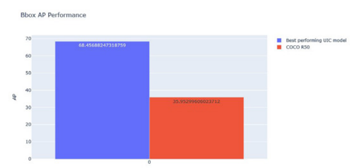



**Figure 5:** Comparison of UIW vs COCO R50 bounding boxes AP

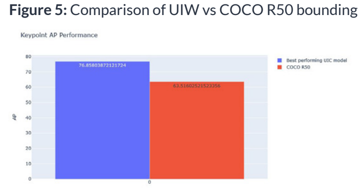



**Figure 5:** Comparison of UIW vs COCO R50 keypoint AP

## Conclusions

- Developed a wheelchair-focused RGB dataset.
- Enable fast improvements in performance on users in wheelchairs in common models for better equity.
- Develop an end to end pipeline for further expansion into the existing UIW dataset.

**Next Steps**
- Expand upon UIW dataset.
- Test other common MMC models.
- Test efficacy of synthetic data and comparisons with the existing UIW.

## Acknowledgements & References

I would like to thank Prof. Yang Zhang and everyone in the HiLab for their support this summer and the National Science Foundation for funding this research. A special thanks to Will Herrera and all the SURP staff for making this program possible.

[1] Lin, Tsung-Yi, Michael Maire, Serge Belongie, Lubomir Bourdev, Ross Girshick, James Hays, Pietro Perona, Deva Ramanan, C. Lawrence Zitnick, and Piotr Dollár. "Microsoft COCO: Common Objects in Context." arXiv, February 20, 2015. https://doi.org/10.48550/arXiv.1405.0312.
[2] Yuxin Wu and Alexander Kirillov and Francisco Massa and Wan-Yen Lo and Ross Girshick. "Detectron2", 2019. https://github.com/facebookresearch/detectron2.
[3] Andres Vasquez, Marina Kollmitz, Andreas Eitel, & Wolfram Burgard (2017). Deep Detection of People and their Mobility Aids for a Hospital Robot. In Proc.~of the IEEE Eur.~Conf.~on Mobile Robotics (ECMR).


## FACULTY ADVISOR
**Yuan Tian**

## DAILY LAB SUPERVISOR
**Tamjid Rahat**

## DEPARTMENT
**Electrical and Computer Engineering**

## XIANGYU WANG

Electrical and Computer Engineering
1st Year, UCLA


# Generic Deobfuscation of Control Flow Flattening

## ABSTRACT


Control Flow Flattening (CFF) is a code obfuscation method used to convolute program logic to impede reverse engineering and hinder analysis. In response to growing worries about software security, our research lab tackles the challenge of untangling CFF by using deobfuscation techniques and partial evaluation. We have two main goals: first, to explain the intricacies of CFF-obfuscated code and its effects on program understanding, and second, to engineer deobfuscation strategies to undo the obfuscation, only restoring meaningful parts of the program for better analysis.

We used OLLVM (Obfuscator-LLVM) to perform CFF obfuscation on programs that we wrote ourselves to better analyze the method and develop how to deobfuscate it. We wrote scripts using the reverse engineering tool Ghidra to analyze the program and its Control Flow Graph (CFG). Our method combines existing deobfuscation methods with partial evaluation, which lets us dynamically unfold the twisted code and rebuild a clearer version of the original. Our early research results are promising, showing that this approach can make sense of CFF-obfuscated programs. This work matters because it helps both defenders and attackers in the world of software security. Deobfuscating Control Flow Flattening not only helps security experts analyze and safeguard software, but it also gives insights to those studying code, finding weaknesses, and dissecting malicious software. Our lab aims to create stronger safeguards for software and better ways to handle the constantly changing tactics of code obfuscation in the realm of cybersecurity.

# Generic Deobfuscation of Control Flow Flattening

Xiangyu Wang, Tamjid Al Rahat, Prof. Yuan Tian

Department of Electrical and Computer Engineering, University of California – Los Angeles

UCLA Samueli School of Engineering

SUMMER UNDERGRADUATE RESEARCH PROGRAM

## Introduction

Deobfuscation techniques encompass a spectrum of approaches, from static and dynamic analysis to machine learning-based methods, enabling a more comprehensive understanding of obfuscated code.

Notable gaps in knowledge still exist in the realm of deobfuscation, particularly concerning highly resilient obfuscation schemes, automated detection of emerging obfuscation tactics, and the development of universally effective tools for tackling diverse forms of code obfuscation.

## Objective

Explain the intricacies of CFF-obfuscated code and its effects on program understanding. Use deobfuscation strategies to restore meaningful parts of the program for better analysis.

## Background/Principles

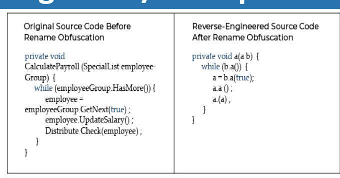



**Figure 1:** Example of a simple obfuscation technique called Renaming, just changing the name of meaningful variables/classes/functions/ect into meaningless strings.

**Obfuscation** – technique to make codes more complex and harder to understand

**Deobfuscation** – reversing obfuscation and attempt to restore the original purpose of the code.

**Obfuscator-Low Level Virtual Machine (OLLVM)** – Tool to perform different Obfuscation methods including CFF on program.

**Control Flow Graph (CFG)**– Diagram of program flow

**Control Flow Flattening (CFF)** – Mixes CFG to confuse analysis. Make the actual code path difficult to trace

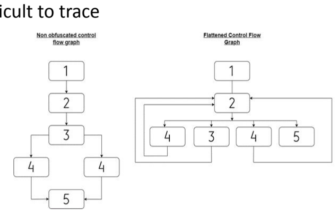



**Figure 2:** Example of CFF obfuscation, left is original CFG and left is after CFF, the overall flow is turned into a huge loop and harder to trace

## Hypothesis

- Using deobfuscation techniques with Ghidra tool and partial evaluation should be able to provide a simpler CFG for malware detection and code analysis.
- Should have less limitations since we are using Ghidra scripts.

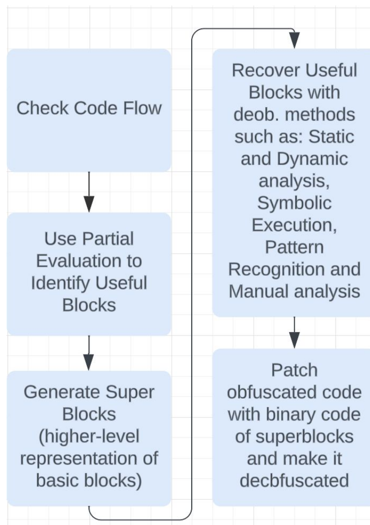
## Methods

Write our own program to obfuscate and analyze

Use OLLVM tool to perform CFF on the program

Put executable binary file into Ghidra tool for reverse engineering.

Write Ghidra scripts to extract the important information from the code and perform deobfuscation analysis.



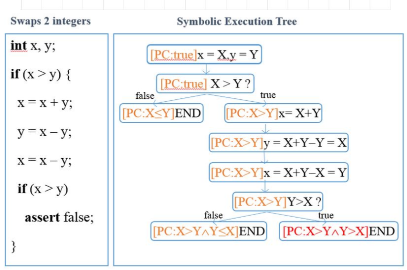




**Figure 3:** Symbolic Execution - analyzing a program to determine what inputs cause each part of a program to execute

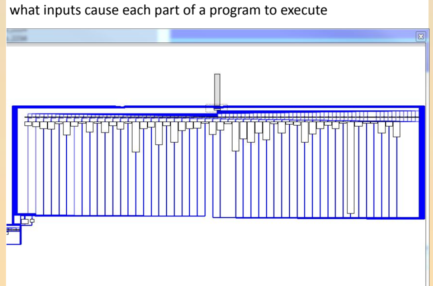


**Figure 4:** CFG of code after CFF

## Results

We were able to install and compile a obfuscator and use it to obfuscate code for analysis

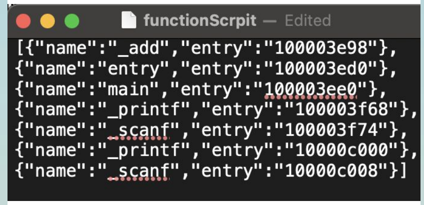



**Figure 5:** We wrote a Ghidra Script that's able to extract and export all of the function names and its entry, we can use this idea to start developing scripts that can analyse the program for us

## Conclusion

We are trying to create a program that can deobfuscate CFF codes with partial evaluation. We try to complete a simplified version first. Understanding the compilation is irreversible, so we just need to understand what does it do. We were able to get out own codes obfuscated with CFF, and we were able to analyse and extract information from it.

## Future Scope

In the future, we intend to conceive a comprehensive software program capable of proficiently deobfuscating CFFed code extracted from applications. This program will subsequently be integrated within the Google Play Store ecosystem, serving as an invaluable tool for code analysts tasked with deciphering obfuscated code structures.

android Google Play

## References

Dong, Weiyu, et al. "CaDeCFF: Compiler-Agnostic Deobfuscator of Control Flow Flattening." Proceedings of the 13th Asia-Pacific Symposium on Internetware. 2022.

Zhang, Xiaolu, et al. "Android application forensics: A survey of obfuscation, obfuscation detection and deobfuscation techniques and their impact on investigations." Forensic Science International: Digital Investigation 39 (2021): 301285.

## Acknowledgements

This work was funded by the National Science Foundation through UCLA Summer Undergraduate Research Program and UCLA Fast Track to Success Program in UCLA Electrical and Computer Engineering department. Also Thanks to Prof. Yuan Tian and lab supervisor Tamjid Rahat for providing guidance and mentorship throughout the research. Special thanks to Will Herrera and all SURP staff for organizing the program.


## ZIXIANG (JERRY) JI

Computer Science
1st Year, UCLA

**FACULTY ADVISOR**
Christina Fragouli

**DAILY LAB SUPERVISOR**
Osama Hanna

**DEPARTMENT**
Electrical and Computer Engineering


# Reduction from Contextual to Linear Bandit

## ABSTRACT


Contextual linear bandits is an important problem with diverse applications in online advertising, recommendation systems, and healthcare interventions. In contextual bandits, a learner interacts with an environment by sequentially selecting actions in the changing action sets based on the context and receiving rewards accordingly, which are measured by the dot product of the chosen action and an unknown parameter. The objective is to minimize the regret, which is defined as the cumulative difference between the rewards obtained by the learner's chosen actions and the optimal rewards in hindsight. The problem is considered more challenging than linear bandits, where the action set remains fixed in each iteration. As a result, more simpler and efficient algorithms have been developed for linear bandits compared to contextual bandits.

In previous work by the lab, where the main contributor was my graduate student advisor, Osama Hanna, a reduction algorithm was proposed to address the contextual linear bandit problem by reducing it to a linear bandit problem. This reduction algorithm enables all developed and future algorithms for linear bandits to solve contextual bandits. However, the reduction framework was only proved theoretically, with its practicality unknown. In this project, we implement the reduction framework from scratch and compare its performance with conventional contextual bandit algorithms. By applying the reduction, we attain similar regret for contextual bandits, as state-of-the-art schemes, using simpler algorithms. Furthermore, we observe that the regret obtained from the reduction method appears to be even lower than conventional approaches for a number of instances.

Samueli Research Scholars

# Reduction from Contextual to Linear Bandit

Zixiang (Jerry) Ji, Osama Hanna, Christina Fragouli

Algorithmic Research in Networked Information (ARNI)
Department of Electrical and Computer Engineering - UCLA

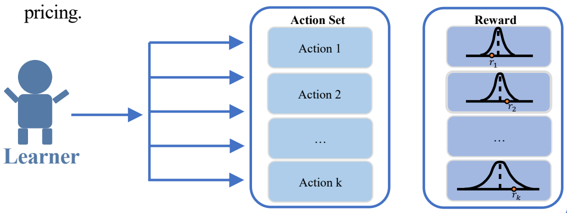
UCLA Samueli School of Engineering
SUMMER UNDERGRADUATE RESEARCH PROGRAM

NSF Grant:
#2007714
#2221871

## Background

### Bandit Problem

- The bandit problem is a foundational concept in decision theory and optimization, with direct relevance to reinforcement learning and machine learning.
- The multi-arm bandit framework proves to be a versatile tool for addressing a multitude of applications: personalized recommendation systems, clinical trial optimization, online advertising, and dynamic pricing.




## Introduction

### Linear Bandit Problem Setup:

- Choose an action from a range of options
- Receive reward based on dot product with the unknown parameter:

$$r_t = \langle a, \theta_\star \rangle + \eta_t$$

- Aim to maximize reward over a sequence of trials

### Contextual linear bandit v.s. Linear Bandit:

- Contextual bandit considers changing dynamic action sets, where the available actions may change over time based on the context.
- Contextual linear bandit is more challenging than the linear bandit in multiple setups
- Therefore, more simpler algorithms have been developed for resolving linear bandit.

## Overview

### Project Overview:

In this section, we simplify the contextual bandit problem by converting it into a linear bandit framework. This allows us to utilize established linear bandit algorithms effectively, providing a practical approach to handling contextual bandit scenarios.

### Evaluation and Comparison:

Our performance evaluation metric is the **cumulative regret**, which quantifies the difference between the rewards obtained and optimal rewards (i.e., achieved by selecting best actions):

$$R_T = \sum_{t=1}^{T} (\max_{a \in \mathcal{A}_t} \mu_a - \mu_{a_t})$$

## Method

### Reduction Algorithm:

#### Definition:

- If the context distribution is known, the fixed action set $\mathcal{X}$ can be defined as the following:

$$\mathcal{X} = \{g(\theta) | \theta \in \Theta\}, \text{ where } g(\theta) = \mathbb{E}_{\mathcal{A}_t \sim \mathcal{D}}[\operatorname{argmax}_{a \in \mathcal{A}_t} \langle a, \theta \rangle | \theta]$$

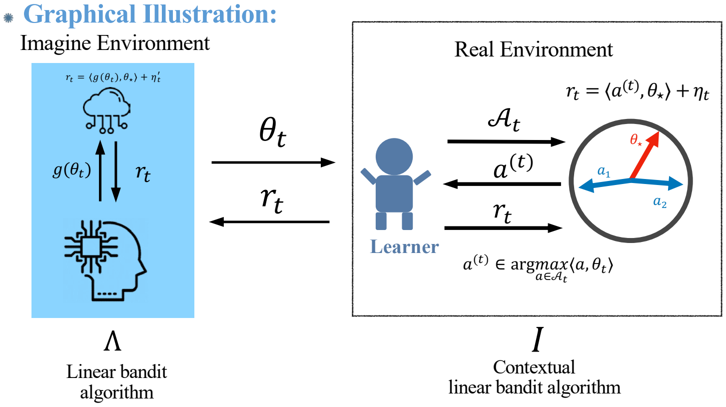
#### Learner-Environment Interaction:

- Learner pulls an arm $g(\theta_t) \in \mathcal{X}$ according to the history observed in the linear bandit algorithm. The action is never actually played. Instead, learner uses the associated $\theta_t$ as the current estimate of $\theta^*$.
- Learner observes $A_t$ and plays the action $a_t = \operatorname{argmax}_{a \in \mathcal{A}_t} \langle a, \theta_t \rangle$
- The learner receives reward $r_t$ and employs efficient linear bandit algorithms to optimize its future action selections.

#### Learner's Objective:

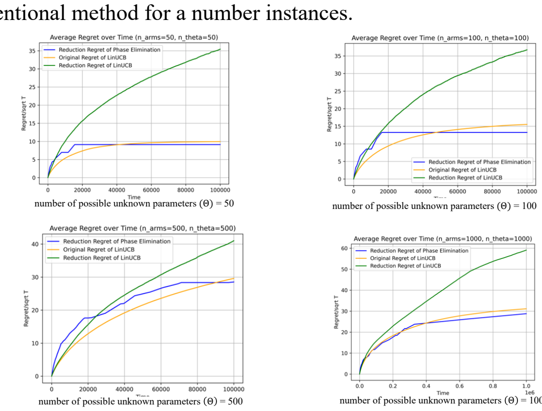
- Iteratively improve action selection, maximize cumulative rewards, and minimize regret.

#### Graphical Illustration:




## Results

As shown in the plot of three methods over 20 runs of experiments, the regret obtained from the reduction method appears to be even lower than the conventional method for a number instances.




If you would like to find out more about the
UCLA Samueli Summer Undergraduate Research Program,
please contact our team at surp@seas.ucla.edu or visit our website at
www.seasoasa.ucla.edu/surp.

seasoasa.ucla.edu/undergraduate-research-program/

Samueli Engineering Undergraduate Research and Internship Program

UCLA Samueli Undergraduate Internship Program

UCLA Engineering Internship and Research Program

samueliurpuip

UCLA Samueli
School of Engineering

UCLA Samueli
School of Engineering

UCLA Samueli Summer Undergraduate Research Program
Office of Academic and Student Affairs
UCLA Samueli School of Engineering
420 Westwood Plaza, 6288 Boelter Hall, Los Angeles, CA, 90095-1600
310.825.9478 | www.samueli.ucla.edu