Multi-Messenger Astrophysics

Michael Coughlin (UMN)
MMA Coordinator

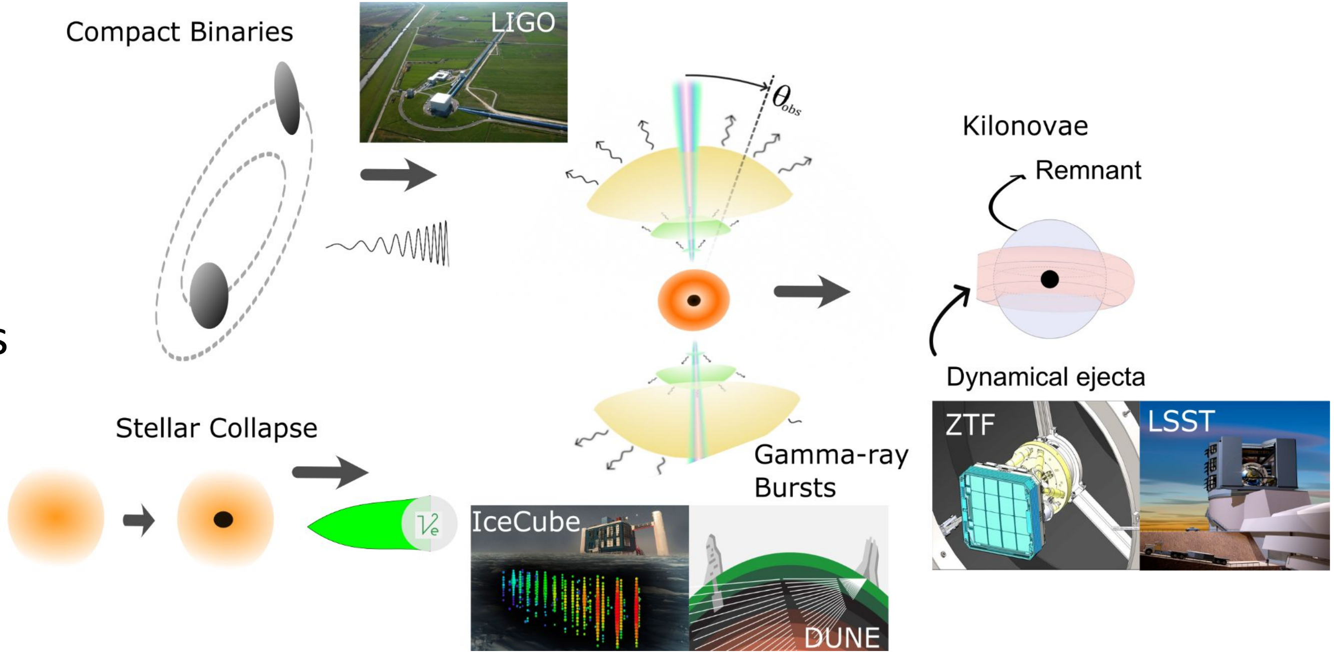
SB Meeting
September 28 2023

OAC-2117997

https://a3d3.ai/

Multi-messenger Astrophysics: Overview

- Our Mission: Develop and deploy software within astronomical facilities
- Our Vision: To use this software to enable discovery



MMA: Research Focus

Accelerated AI Algorithms for Data-Driven Discovery

The **Multi-Messenger Astrophysics (MMA)** pillar is currently revolved around preparation for ongoing and forthcoming observing runs by the astronomical facilities.

In particular, we are focused on development to improve and characterize transient source detection(s) by **gravitational-wave detectors** such as Advanced LIGO, Virgo and KAGRA.

Associated with such transient source detections/candidates, we are contributing to the preparations for the follow-up by **optical facilities** such as the Zwicky Transient Facility of these events, which began May 2023.

An additional direction is development of low-latency triggering and pointing algorithms for supernova burst **neutrino direction**, in the context of DUNE, the SuperNova Early Warning System as well as reconstruction algorithms in IceCube.

Overview of the MMA Team

a3d3 Accelerated AI Algorithms for Data-Driven Discovery

- Monthly meetings across the focus groups, rotating amongst them highlighting the problems and technologies being built to address them.
    - February 20: 10 minute presentations by members from each pillar, group discussion of strategies to promote more cross-pillar discussion
    - March 20: Presentation from Alec Gunny on ML4GW
    - April 24: Presentation from Niharika Sravan on observing scenarios plus follow-up optimization
    - May 22: Presentation from Benedikt Riedel on IceCube machine learning initiatives
    - June 26: Presentation from Ethan Marx on gravitational-wave detection and parameter estimation
    - August 21: Presentation from Joshua Queen on DUNE source reconstruction
    - September 25: Presentation from Ryan Raikman on GW anomaly detection

- Weekly meetings for each focus group
    - Including meetings open to the broader community

- Focus on public data sets
    - Gravitational-wave noise regression sample set (DeepClean paper)
    - Gravitational-wave plus optical observing scenarios (https://arxiv.org/abs/2306.09234)
    - IceCube kaggle competition

UMN (GWs, Optical): Coughlin, Healy, Muhammed, Barna, Benoit, Omer, Toivonen

MIT (GWs): Katsavounidis, Harris, Gunny, Chatterjee, Govorkova, Marx, Moreno, Raikman

Caltech (Optical): Graham, Jegou du Laz

Drexel (Optical): Sravan, Pletskova

Duke (Neutrinos): Scholberg, Hakenmueller, Queen

WIPAC (Neutrinos): Riedel, Weber, Peterson, Hanson

Overall: 8 faculty / senior members, 7 Post-Docs / Computing Staff, 11 students

Metrics of success

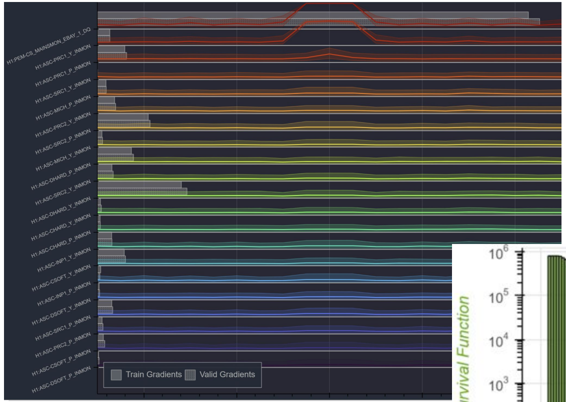
	Gravitational waves	Optical Astronomy	Neutrinos
Number of github repos	9 (4 algorithms, 5 infrastructure)	3	1
Number of research presentations within our collaborations	More than a dozen to LVK subgroups	~ half a dozen to ZTF machine learning group and collaboration meetings	~ 10 with DUNE ~ half a dozen within IceCube
Number of research presentations outside our collaborations	Each algorithm plus infrastructure was presented at APS	Scope was presented at APS, Pythia in multiple ML-focused venues, Observing scenarios at Open LVK-EM telecons	Presented in MMA-focused venues
Number of new contacts within our collaborations	5 (NYCU-Taiwan, AEI, UvA, UWA, Cardiff)	1 (Northwestern)	2 (Fermilab, Columbia)
Number of new contacts outside our collaborations	0	1 (LANL)	0

Accelerated AI Algorithms for Data-Driven Discovery

Gravitational Waves - Overview

Accelerated AI Algorithms for Data-Driven Discovery

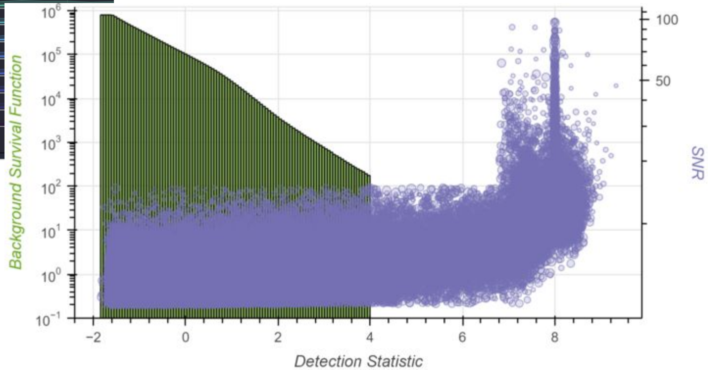
Clean the Data: DeepClean (CNN)



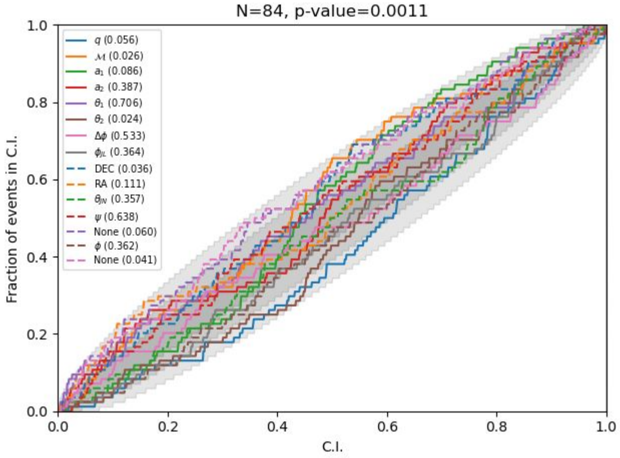
Top-level github work area:
ML4GW

Detect the GWs: aframe (CNN)/GWAK (autoencoders)

~3 faculty, 3 postdocs, 6 grad students, 2 postbac/undergraduates



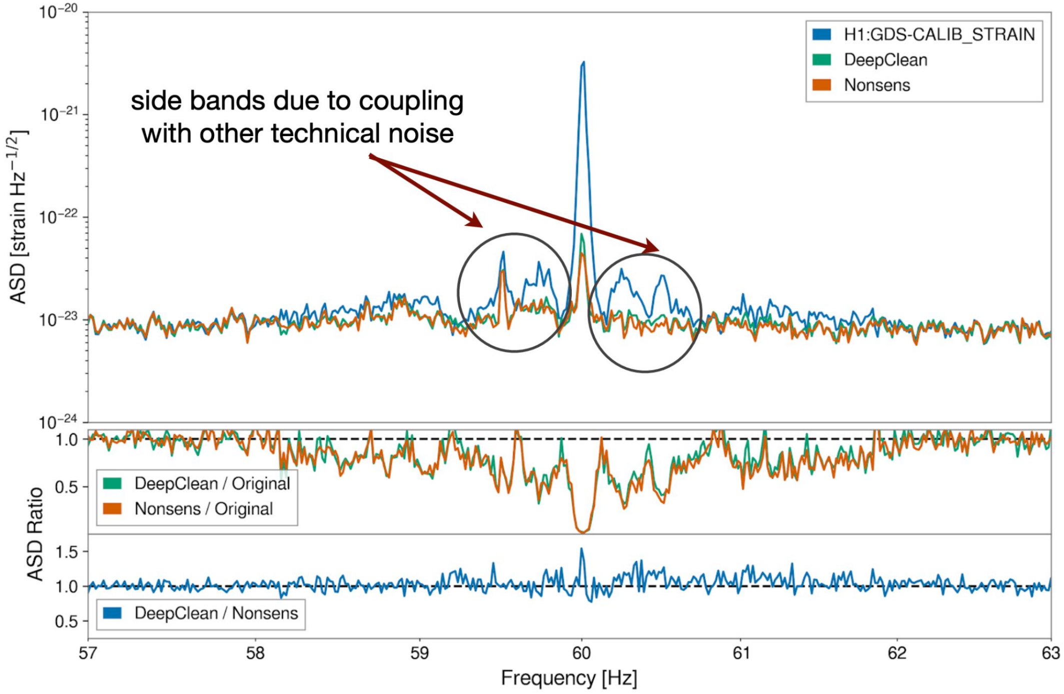
Characterize the GWs: (MAF*)



Main focus: All algorithms use our inference-as-a-service (IaaS) prototype to implement a real-time noise subtraction pipeline (DeepClean), detection (aframe/GWAK), and parameter estimation for use during the fourth observing run (O4) of LIGO-Virgo-KAGRA on dedicated hardware at the detector sites.

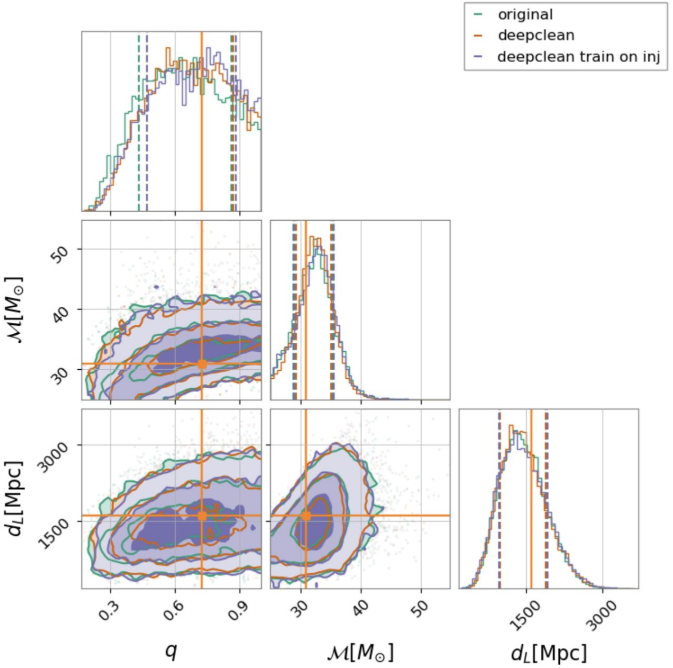
MAF: Masked Autoregressive FLow

Gravitational Waves - DeepClean

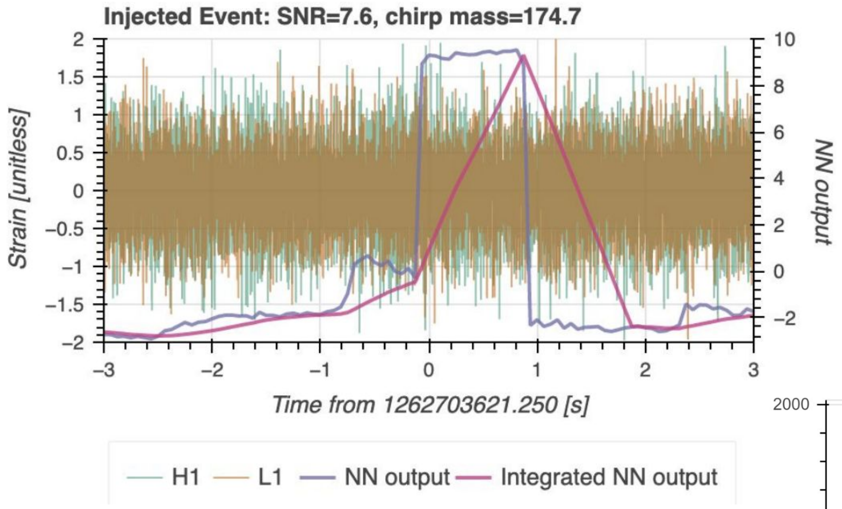


Demonstrated non-linear subtraction on 60 Hz power lines and sidebands!

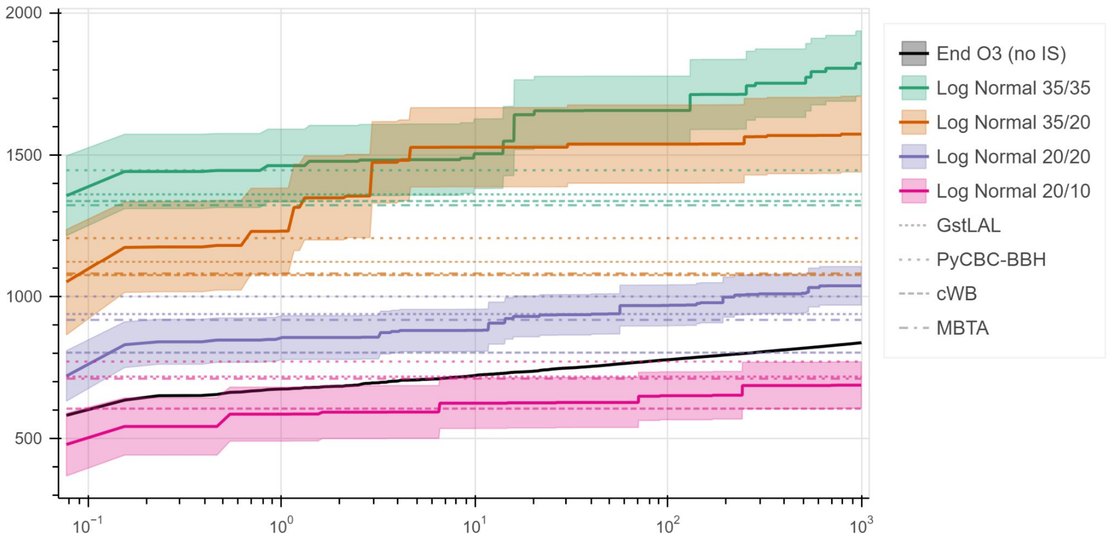
No adverse effects on source reconstruction.



Gravitational Waves - aframe



- Model is run on timeslides of background and background with injections
- Integrated NN output is used as a ranking statistic to estimate FAR and sensitive distance

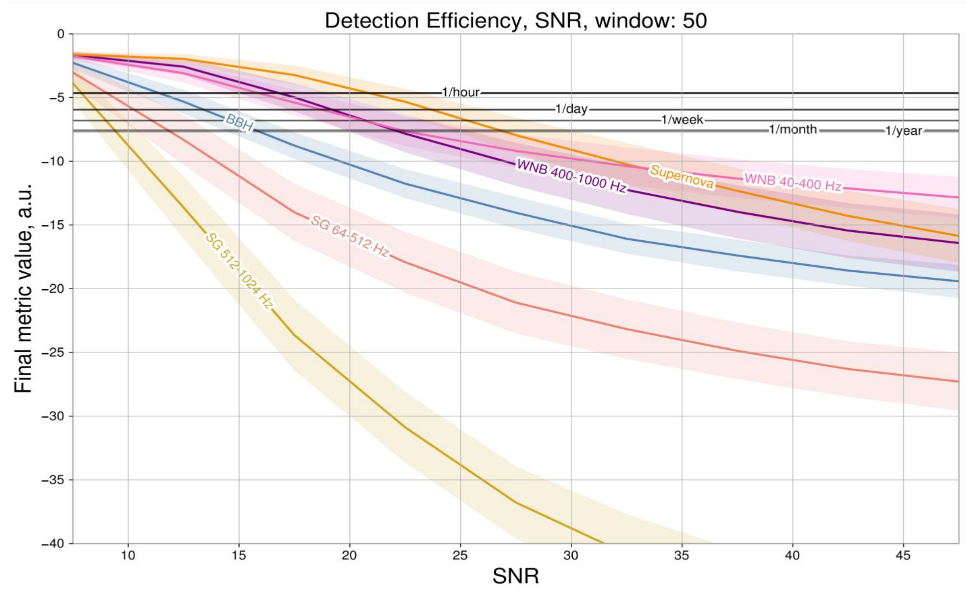


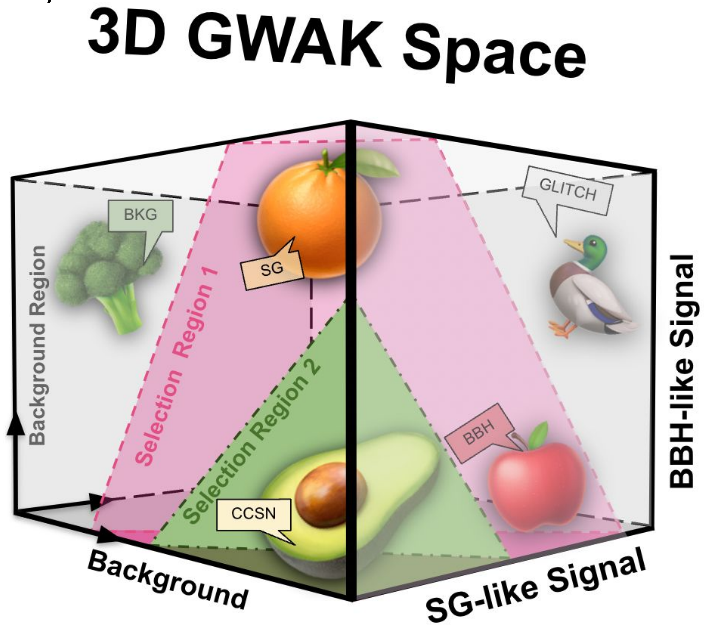
Comparable sensitivity to existing matched filtering algorithms, but faster and with many ideas for sensitivity improvements to come!

Gravitational Waves - GWAK

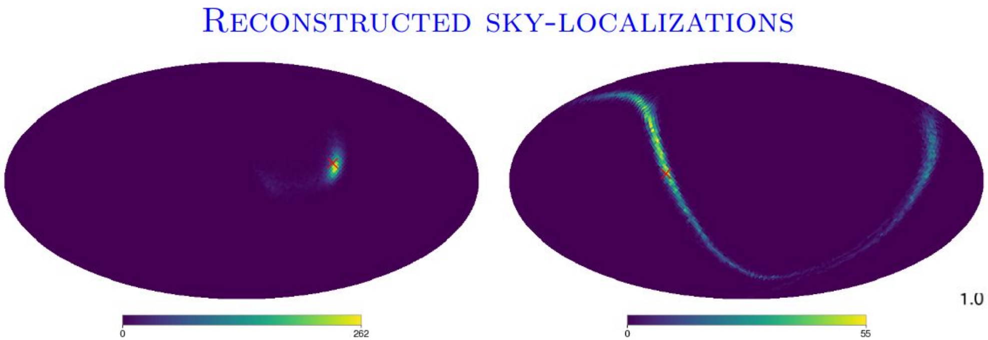
Gravitational-Wave Anomalous Knowledge with Recurrent Autoencoders

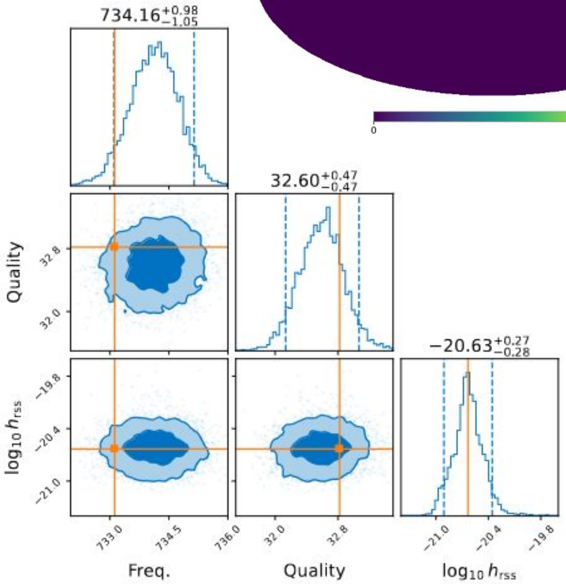
- A new **semi-supervised approach** to gravitational-wave **anomaly detection**
- Our method constructs **a 12-dimensional embedded space** (GWAK space) that is cut with a hyperplane
- We produce background timeslides to verify the FAR of the model and test on different signal/anomaly priors (BBH, SG, WNB, CCSN)
- We find that the **GWAK** method provides **good discrimination power** over correlation metrics **for anomalous signals** that it was not trained on (WNB, Supernova)



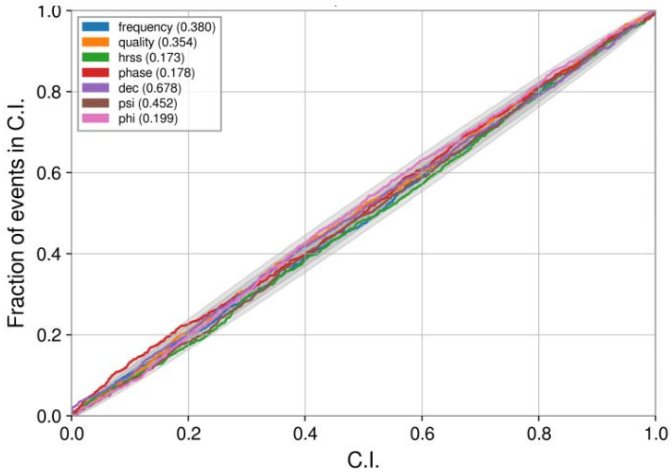


Gravitational Waves - Parameter Estimation





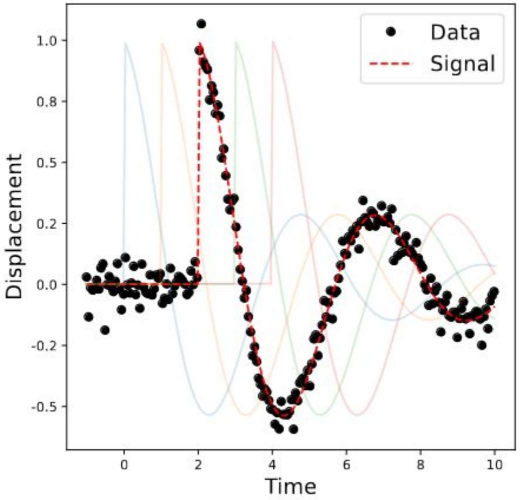
Parameter estimation consistent with conventional samplers, but with hundreds of samples generated per second.



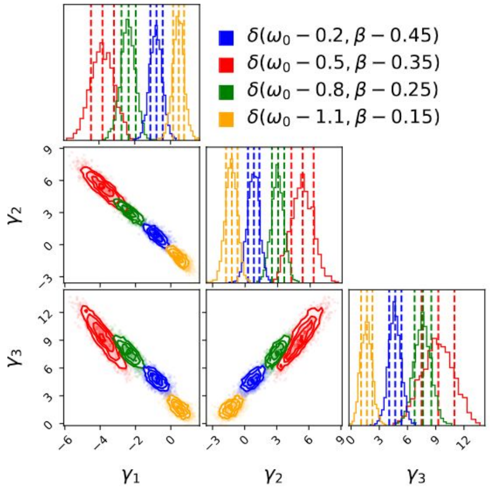
Probability-probability plots show posteriors consistent over a large sample across parameters.

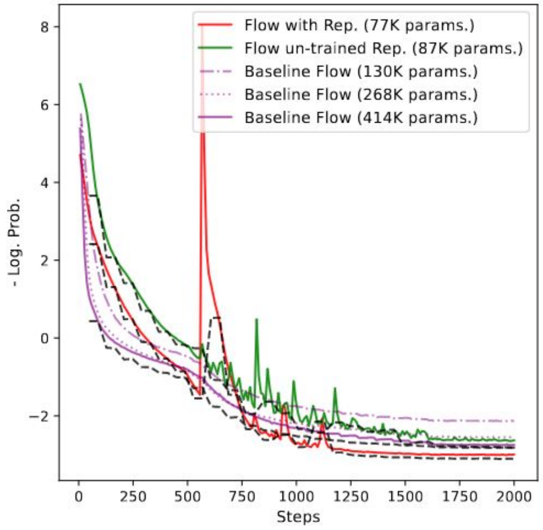
Gravitational Waves - Parameter Estimation

Proof of concept demonstration of incorporating symmetries and optimizing likelihood free inference



Similarity embedding for time-translation





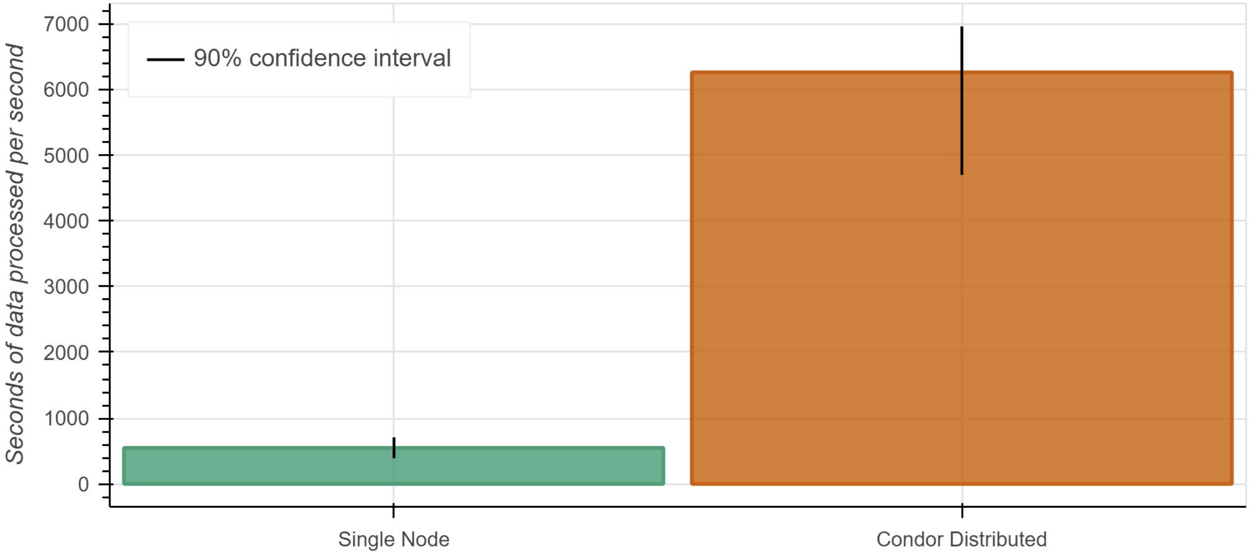
Same performance at lower number of parameters. Currently exploring avenues of using this in GW and KNe parameter estimation.

Gravitational Waves – Community Impact

Has brought together research groups pursuing similar problems within the LVK collaborations and led to active participation from a broader community beyond A3D3 members, code and infrastructure sharing and common development.

A3D3-led software (HERMES) has already assisted and accelerated over ten-fold the ability to process background data for a ML-based transient detection algorithm developed by the Cardiff group in the LVK.

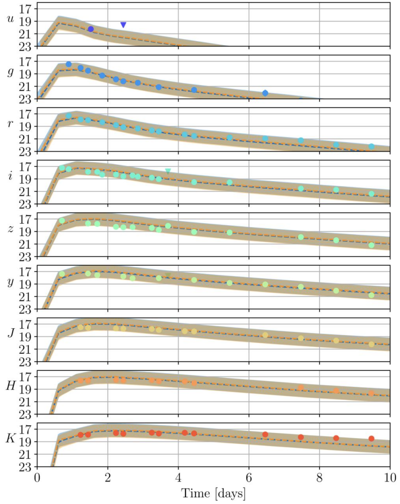
Additionally, A3D3-led software (DeepClean) is assisting and contributing to algorithmic reviews and cross-validation of similar software within the LVK.



Optical Astronomy - Overview

Accelerated AI Algorithms for Data-Driven Discovery

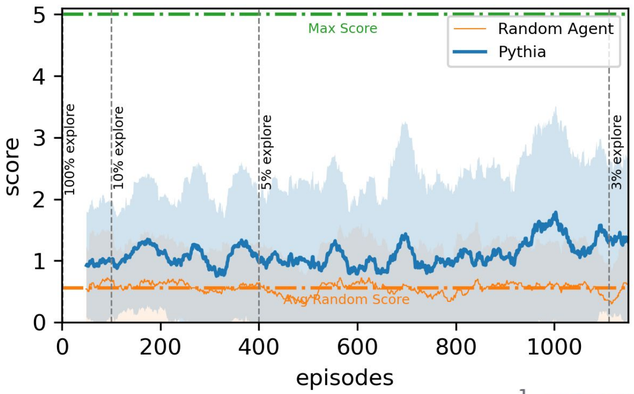
Simulate Observations: NMMA (emulator)



Github work areas:

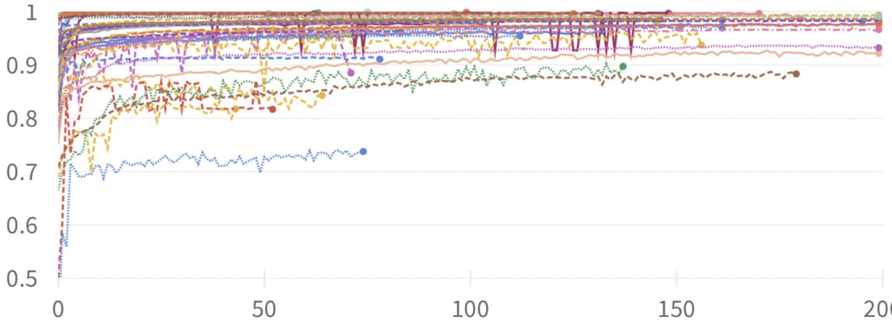
NMMA SCOPE Pythia

Optimize Observations: Pythia (RL)



~4 faculty, 3 postdocs, 5 grad students, 3 postbac/undergraduat

Classify the sources: Scope (CNN)



Main focus: Deploy ML algorithms throughout the observation preparation and follow-up for source identification and characterization

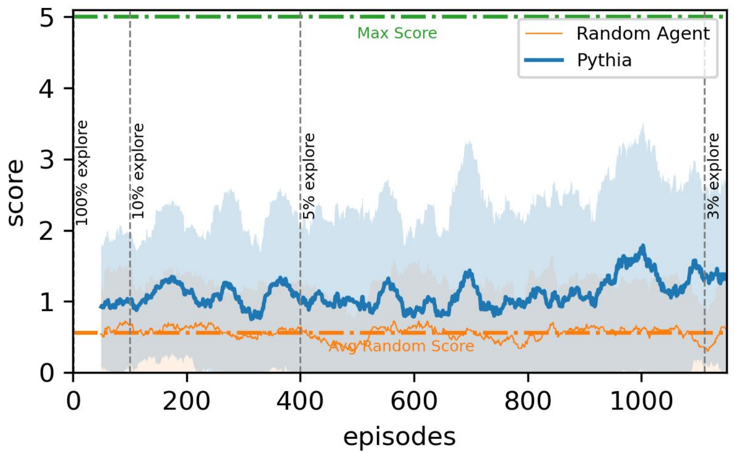
Optical Astronomy – Pythia

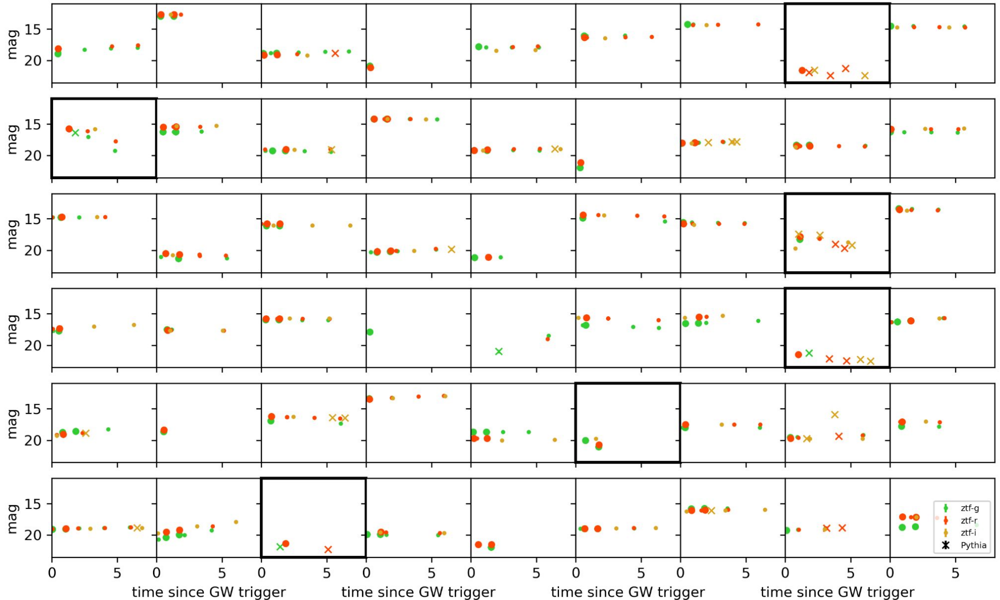
SARSA agent (linear VFA and TD(0) target) to maximize additional photometric follow-up in ZTF g, r, or i allocated to KNe vs. contaminants

- Agent 3x better than random. 80% of KNe receive at least one follow-up photometry
- Overall prefers i filter on days 1-2 and r filter after
- Confuses KNe with GRB afterglows on days 2-4 and SNe on days 4-6
- Human benchmarking in progress. Paper in prep. Code: https://github.com/niharika-sravan/Pythia

Transitioning to deep Q networks:

- Developing environment in Gymnasium to attach to DQN agent in tf-agents.





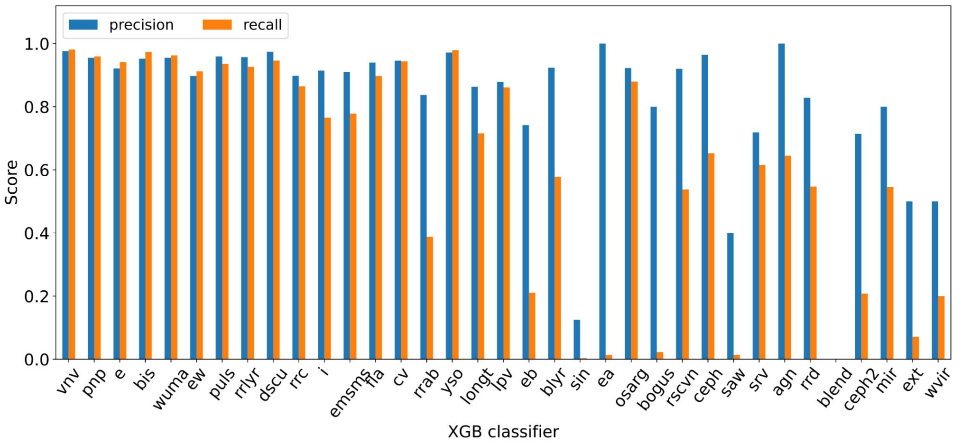
Optical Astronomy – scope

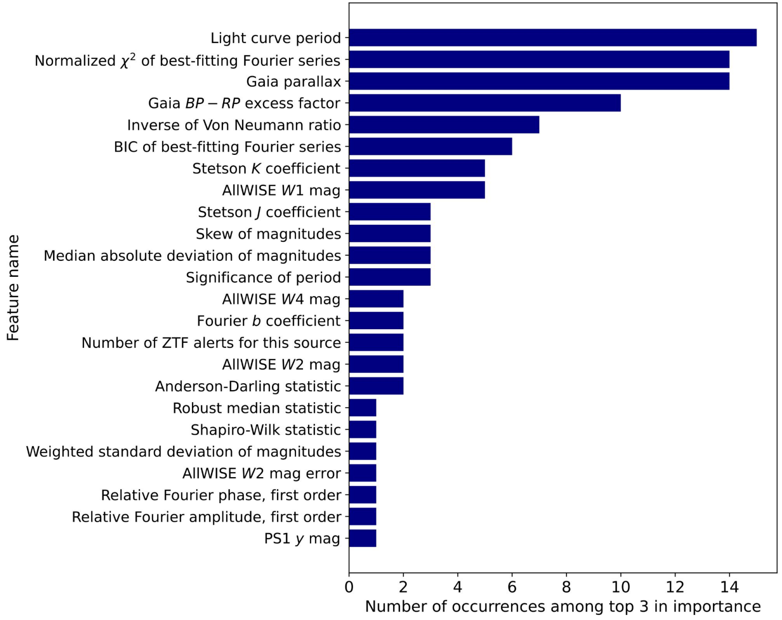
ZTF Source Classification Project (SCoPe):

Provides ML source classifications to assist GW counterpart follow-up, variable star studies

- Light curve feature generation on ZTF DR16 runs c ACCESS (SDSC Expanse)
  - Small scale: GCN sources of interest
  - Large scale: many/all ZTF fields
- Neural network/XGBoost training is complete
  - Achieving high precision/recall for many of the 35 binary classifiers per algorithm
  - Studying the most important features to streamline future classification
- Talk at APS April Meeting 2023; paper describing upcoming catalog of ZTF ML classifications for ~10^8 light curves is nearing submission

SCoPe code publicly available here:
https://github.com/ZwickyTransientFacility/scope





Optical Astronomy – Community Impact

Accelerated AI Algorithms for Data-Driven Discovery

Observing Scenarios:

Publicly available on zenodo: https://zenodo.org/record/7026209
1084 views, 87 downloads

Gravitational-wave follow-up:

Reported by GCN:

1. LIGO/Virgo/KAGRA S230627c: Zwicky Transient Facility observations and candidates
2. LIGO/Virgo/KAGRA S230602ap: Zwicky Transient Facility observations
3. LIGO/Virgo/KAGRA S230529ay: Zwicky Transient Facility observations
4. LIGO/Virgo/KAGRA S230528a: Zwicky Transient Facility observations
5. LIGO/Virgo/KAGRA S230527ch: Zwicky Transient Facility observations

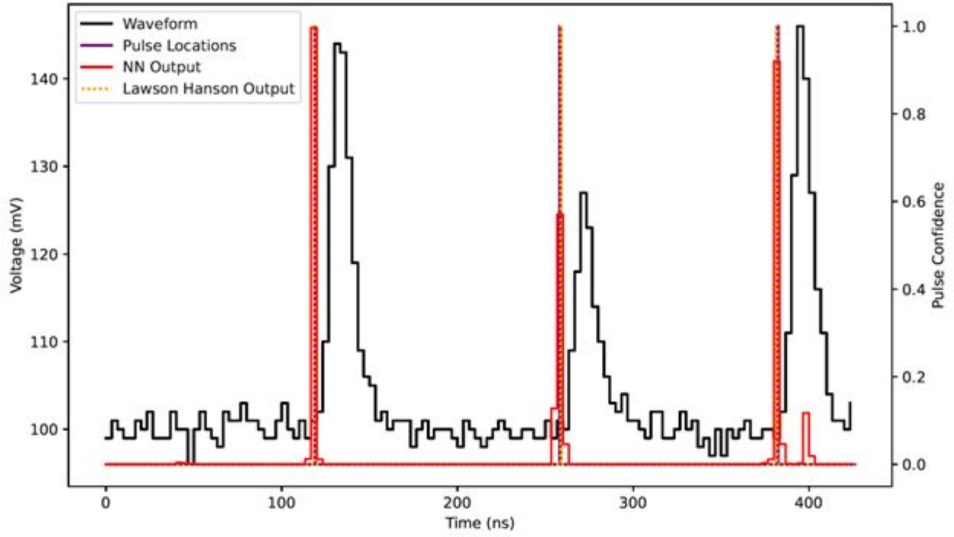
Observations publicly reported on Treasure Map and candidates on TNS
Data from MSIP gravitational-wave skymap weighted fields immediately public

Expansion of transient ML capabilities:

Built in generic inclusion of ML classification into ZTF filtering software stack (Kowalski): Now being used to automate bright transient identification for the “Bright Transient Survey” with a multi-input convolutional neural network – **First automated trigger in our program achieved yesterday**

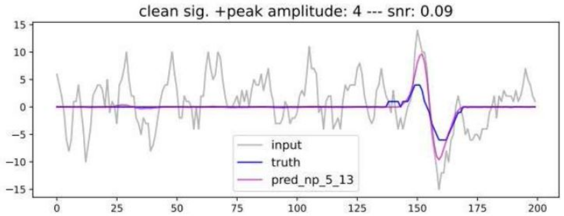
Neutrinos - Overview

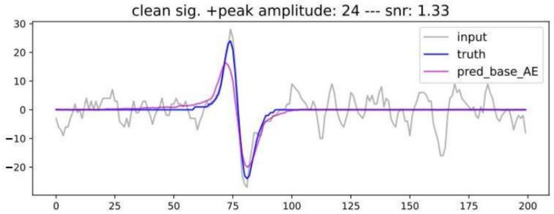
PMT Voltage Picking (CNN)



~2 faculty, 2 postdocs, 2 grad students, 2 postbac/undergraduates

Supernova Reconstruction (1DCNN autoencoder + pointing)





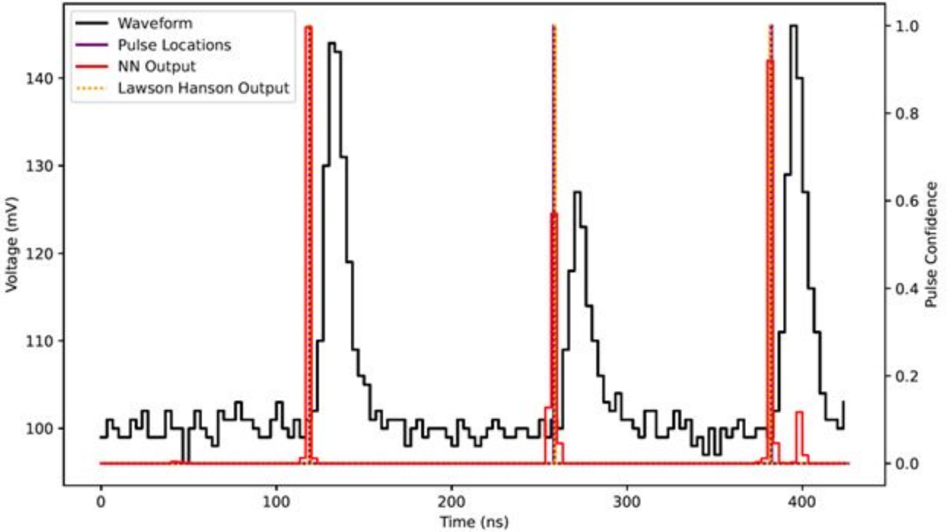
Main focus: Porting existing algorithms to GPUs and FPGAs for the purpose of detection and localization reconstruction.

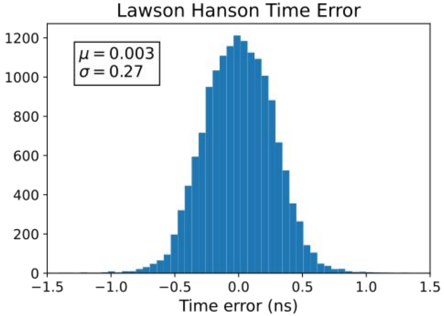
See: See Pan's Talk in Hardware-Algorithm Co-Development

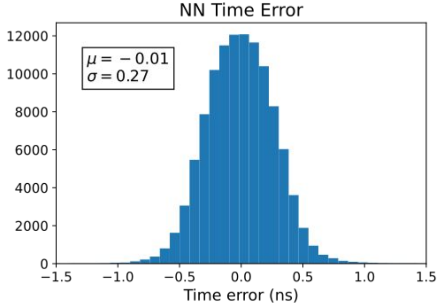
Neutrinos - Activities

IceCube

- Goal: Compare neural networks to traditional algorithms in process PMT voltage waveforms in GPU
- Developed a 1D CNN for finding populated voltage waveform bins and a small neural network for searching populated bins for photon hits - Timing resolution the same
- We have ported the existing waveform deconvolution algorithm from the IceCube Upgrade microcontroller design to a more general Xilinx implementation to allow using the algorithm at the IceCube counting house at the South Pole.
- Goal: Use high level synthesis tools to deploy wavedeform algorithm on accelerator card (single copy successful, but parallelized copy did not yield the acceleration warrented by FPGA)

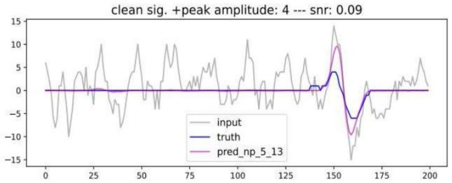


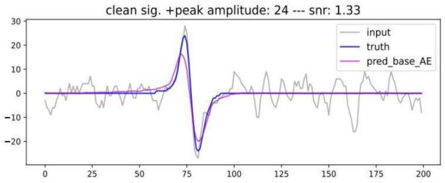


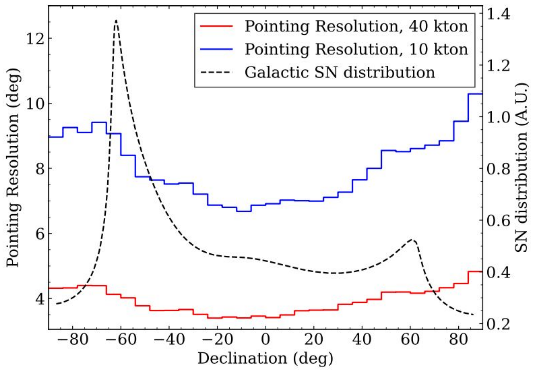


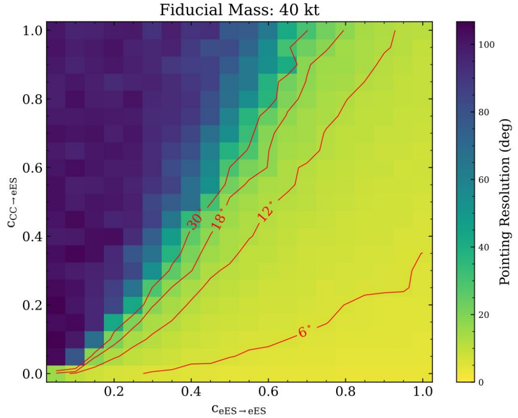
Neutrinos - Activities

- Overall goal: DUNE low-latency pointing capability (fast reconstruction)
- Non-ML pointing proof-of-concept so far: paper in DUNE collaboration review
- Depends on “channel tagging” (event classification)-> ML effort
- Also: 1DCNN autoencoder for hit ROI finding in TPC induction plane waveforms
- Underway: : implement in in-situ compute setup
- PINNS-AIHEP effort: ICEBERG data @ FNAL this summer (Janina H /Josh Q)
- **New post-bacc Lucie Afko getting started**
- **Supernova pointing Hack Days @ FNAL organized by Janina**









Van Tha bik Lian (past post-bac), Lucie Afko (new postbac), Janina Hakenmueller, James Shen (now grad at Penn), Josh Queen, Kate Scholberg

Neutrinos - Community Impact

DUNE:

More than ten presentations within DUNE (and growing connections with a number of other groups including Michael Wang's at Fermilab and Georgia Karagiorgi's at Columbia) on the applications of ML within the collaboration on the low-latency supernova pointing development and its unique potential for MMA contributions in the event of a core-collapse supernova.

IceCube:

About half a dozen presentations to IceCube on the implementation of the waveform deconvolution algorithm in accelerators and microcontrollers, which is essential for IceCube Upgrade and IceCube Gen2., given that reducing IceCube's power footprint at the South Pole is on the critical path for these experiments.

Plans For Future Years

Gravitational waves:

- Demonstrate the noise subtraction online on O4 data
- Run the detection pipeline, both offline and online, during the observing run; seek to deploy in production and publish results.

Optical:

- Perform the first end-to-end follow-up of a gravitational-wave event and publish it.
- Finish the data release and first training of the optimization algorithm
- Finalize a catalog of variable ZTF sources to be published.

Neutrinos:

- Work towards the development of real-time algorithms for supernova detection within liquid argon systems, and understand the effectiveness of deploying our models in real-time.
- Extend the complexity of the models we have run as-a-service with GPUs on ProtoDUNE data, and establish this reconstruction chain as default in future neutrino event reconstruction workflows.

Accelerated AI Algorithms for Data-Driven Discovery

Thank you!

Neutrinos - Overview

DUNE/SNEWS/COHERENT

Faculty: Kate Scholberg
Postdoc: Janina Hakenmüller
Grad student: Joshua Queen
Post-bacc: Van Tha Bik Lian. New in August: Lucie Afko
Specific projects now underway:

- new fast SNB pointing working group led by Hakenmüller (weekly meetings)
  - First paper draft indicating DUNE pointing capability (starting collab review)
  - Starting implementation on in-situ computing setup
  - ICEBERG (prototype neutrino detector) data this summer
- VTB Lian working on denoising raw DUNE waveforms w/autoencoder

Group meetings on zoom: Fridays 4 pm Eastern Time + other collaboration-specific meetings

IceCube

- Working on accelerating the MMA pipeline by using accelerators and ML methods to reduce time to result
- Largest compute and timesink is waveform deconvolution
- Need to test deployments with GPU and FPGA-based “traditional” algorithm compared to ML-based algorithm

Impact

On the development of the principal discipline(s) of the project:

All three pillars are very well integrated into and driving the growth of their respective communities

On other disciplines

With the gravitational-wave work, we have curated and provided a data set to test novel regression algorithms against, and will be doing something similar for the detection work. On the optical side, we are creating a novel fast transient focused data set with Bayesian inference posteriors provided; we are also enabling the observing scenarios for the next gravitational-wave observing run. All of these provide unique data sets for others to test out algorithms with.

On the development of human resources

The projects all have significant graduate student and undergraduate involvement in a variety of forms, with post-docs and graduate students the point of contacts for a number of deliverables.

On the development of educational experiences

The data sets being curated will be used in the ZTF Summer School and in a variety of classes taught by the A3D3 professors.

Institutional resources that form infrastructure
N/A

Outcomes, Achievements, and Objectives

Gravitational waves

- Bringing in other data analysis pipelines (e.g., detection/classification, parameter estimation and interferometer controls) within the International GW network (IGWN) into the inference as a service framework
- Established connections with the NYCU-Taiwan, AEI, UvA, UWA and Cardiff groups (Accelerating Physics with ML@MIT very helpful for this)
- Goal: Running noise subtraction online on O4 data (began a couple days ago, pipeline review ongoing); running BBHNet offline over O4a, online during O4b

Optical Astronomy

- Software stack to trigger ZTF demonstrated on early O4
- Goal: finish paper on observing scenarios / follow-up strategy implications / data set

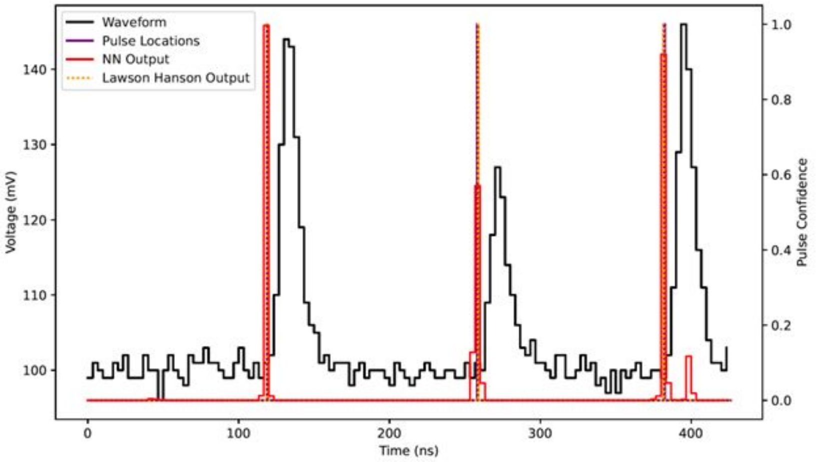
Neutrinos

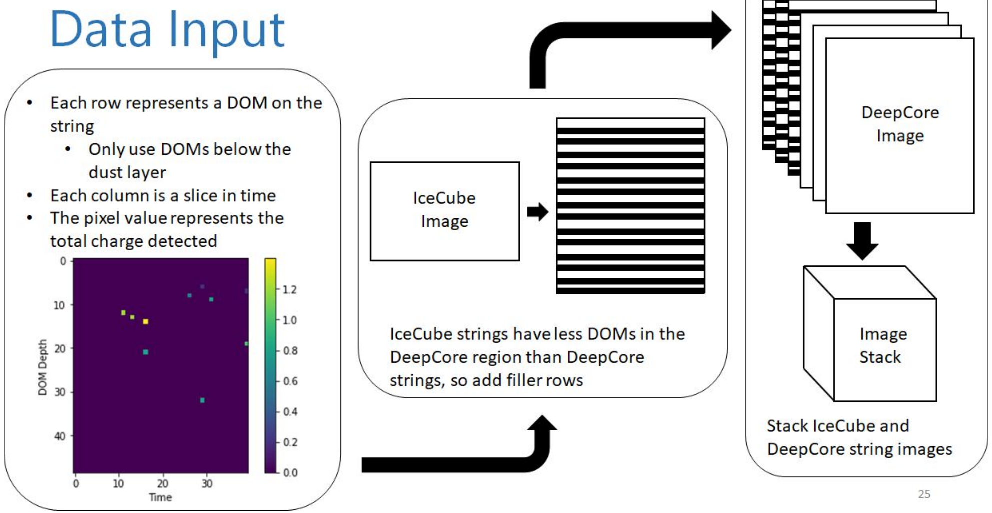
- Developing new fast reconstruction tools for supernova burst pointing in DUNE -> new working group underway**, first sim/reco results;** waveform unfolding in IceCube
- Goal: ML-based channel-tagging for fast pointing reco for DUNE; accelerated unfolding algorithms on the surface for new ML based reconstruction algorithms for IceCube

Neutrinos

IceCube

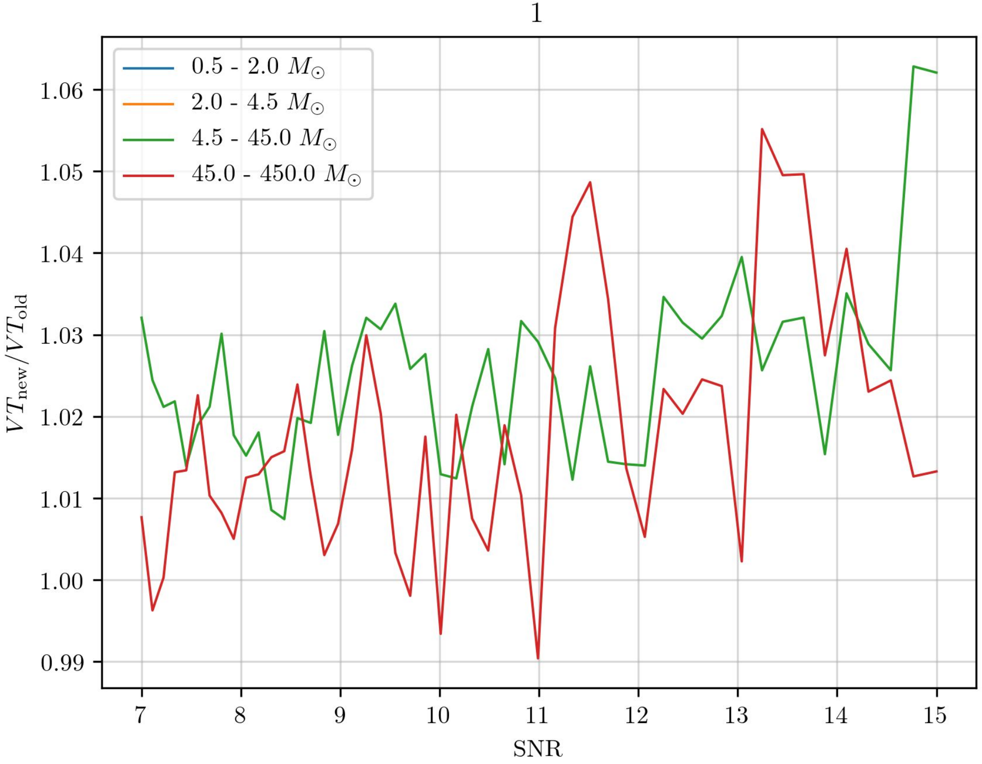
- Started bi-weekly/monthly meetings with Vladimir to talk about HLS issues
- Testing Vitris-based implementation for “standard” algorithm
- Creating CNN for Neutrino oscillation studies
- 1D CNN for waveform deconvolution still development





Gravitational Waves – Denoising work

Improvements in detectability estimated using GstLAL pipeline - preliminary results shows improvements after the subtraction of 60 Hz power-line noise



Neutrinos - Impact

Dune:
The low-latency supernova pointing development is still in progress, but we expect significant impact on DUNE DAQ implementation to enhance its unique potential for MMA contributions in the event of a core-collapse supernova.

IceCube:
The implementation of the waveform deconvolution algorithm in accelerators and microcontrollers is essential for IceCube Upgrade and IceCube Gen2.
Reducing IceCube’s power footprint at the South Pole is on the critical path for these experiments.
Building a faster MMA pipeline will allow faster alert issuing and follow up in the Northern Hemisphere.

* Acceleration of code on different platforms
* We have one grad student and postdoc right now
* Research presentations: ~5
* Outside collaborations: 1 we started talking to NVIDIA