ML challenges in Theia and WBLS

Presented by

Björn Wonsak

on behalf of the THEIA collaboration

NPML, 10th July 2020



Theia Concept/Motivation

- New-generation of large-scale, low-threshold, directional detectors
- Combine advantages of Water-Cherenkov- and liquid scintillatordetectors ASDC Whitepaper October 2014 arxiv:1409.5864

Advantages:

- Scintillation: High light yield
 - \rightarrow Good energy resolution + low threshold
- Cherenkov-light: Non-isotropic emission
 - \rightarrow Directional information
 - + Particle identification by ring-characteristic
- Combined: C/S-ratio
 - \rightarrow Improved particle discrimination

Crucial: Need to be able to separate both light species!

Also important: Low cost & high transparency of water



Number of Cherenkov- and scintillation photons for different particles in LAB

Cherenkov-/Scintillation Light Separation

Three different signatures + new technologies

- Cherenkov-rings → optimize light ratio
 - → Water-Based-Liquid-Scintilator (WBLS)/low light yield scintillator
- Emission time profile
 - \rightarrow fast timing (sensors)/slow scintillator (cocktail)
- Wavelength
 - → filtering/optimized sensors

+ Advanced reconstruction methods







The THEIA Detector

- Large-scale detector (30-100 kton)
- Water-based LS target
- Fast, high-efficiency photon detection with high coverage
- Deep underground (e.g. Homestake)
- Isotope loading (Gd, Te, Li...)
- Flexible! Target, loading, configuration
 - ➡ Broad physics program!

30 kt detector, that would fit 4th DUNE cavern

Concept paper: arXiv:1409.5864

White paper: M. Askins et al., Eur.Phys.J.C 80 (2020) 5, 416, arXiv:1911.03501



10/07/20

Theia Physics Program



Neutrinoless double beta decay







Nucleon decay





S.M. Usman, et al., Scientific Rep. 5, 13945 (2015)



and more ...

Using Other Experiments as R&D Testbeds



Using Other Experiments as R&D Testbeds



How can Theia profit from ML?

Make the most from the data:

- Particle identification
- Cherenkov-Separation
 - Ring counting/analysis
 - \rightarrow Particle identification
- Direction analysis
- Topological reconstruction

Speeding up MC production

 Maybe combine reconstruction & simulation in an invert-able network

Optimizing the detector design

Will not cover this here!

 Requires understanding how data quality and performance are related to detector properties!

How can Theia profit from ML?

Make the most from the data:

- Particle identification
- Cherenkov-Separation
 - Ring counting/analysis
 - \rightarrow Particle identification

- Direction analysis

Many different design options

- \rightarrow Reconstruction algorithms need to be portable
- \rightarrow This is easier for ML

(Once I have architecture, I can train on many different data sets)

in an invert-able network

Optimizing the detector design

Will not cover this here!

 Requires understanding how data quality and performance are related to detector properties!

Particle Identification at MeV Energies

Detector environment:

- JUNO full MC

 \rightarrow 20kt LAB-PPO + bis-MSB (pure LS)

 \rightarrow ~3% Cherenkov-light (2/3 of this is scatter)

total coverage ~80%

~5000 dynode PMTs (σ_{TTS} = 1.27 ns)

~12000 MCP PMTs (σ_{ττs} = 5.1 ns)

→ Much worse TTS than Theia and no WBLS

Assumptions:

- Know vertex from previous reconstruction methods
- Compared three methods:
 - Gatti: Time of Flight (ToF) corrected time spectrum of all hits
 - **ML**: Uses the same data (1D-data)
 - Topological reconstruction (TR): 3D picture of event signature (+ cut based analyses / ML on 3D)

(TR: Optimized for electron events)

Results Particle Identification I

- Discrimination based on long tail of α and proton time spectrum



- ML slightly better than Gatti
- TR not compatible (but also not optimized for this)

Results Particle Identification II

• Discrimination based on topological differences

(additional y, several Compton scattering points, etc.)



- ML best for e+/e- but TR best for e-/γ
- Gatti not compatible (but also not optimized for this)

Results Particle Identification III

- Data-set 1: No TTS, perfect vertex, no DCR
- Data-set 2: Added TTS and realistic vertex
- **Data-set 3:** Added Dark Count Rate (DCR)



Gap between data-set 1 and 2 indicates huge potential of good TTS (good TTS will also affect the vertex resolution)

L. Ludhova et al. arXiv:2007.02687

Machine Learning Example: C-10

- Studied in A. Li et al., arXiv:1812.02906
- Using a Convolutional Neural Network (CNN)
- In KamLAND-like detector (~1ns σ_{T} , 23% QE, 16% coverage)
 - → 62% bkg reduction at 90% signal efficiency 83.5% for JUNO-like coverage and QE, 98% for perfect light collection
 (time delay of ortho-positronium decay not used)
- C-10 is background (bkg) for solar-v and $0v\beta\beta$
- I see similar potential for I-130 & Cs-136 ($0\nu\beta\beta$ bkg)

Conclusion: High granularity and statistics (coverage) are important!



Directional Reconstruction: What Kind of Network do we Use?

· CNN:

- CNN use structured data in Cartesian space
- Converting unstructured and sparse 3D Data (our PMT-positions) can lead to loss of information and quantization problems
- Computation intensive if volume approach or a combination of several views is used to handle 3D point clouds

Graph-Neural-Networks

- Use unsorted nodes + connections (edges)

 \rightarrow All operations need to be invariant against permutation (like max or average)

- Nodes and edges can have features (color, id, ...)
- Nodes usually not equidistant \rightarrow hard to use kernel based filters
- Hard to get power of convolution integrated



 Tx_{1}

X, (

 $\mathbf{X}_{j_{i4}}$

 $\mathbf{X}_{j_{i2}}$

 $\mathbf{x}_{i_{ij}}$

Choice of ML-Tool: PointNet

PointNet:

- Optimized for point clouds
- Can do classification and segmentation
- Each point is a vector (x,y,z) + features (no edges)
- Operating on each point independently
- Subsequently applying a symmetric function to accumulate features (max pooling)
 - \rightarrow Invariant against permutation by using max pooling
 - + not so good in capturing local features
- Robust to various kinds of input corruptions



Choice of ML-Tool: DGCNN

• Dynamic Graph CNN (DGCNN):

Yue Wang et al.: https://arxiv.org/abs/1801.07829

- Independent neural network module
- Can be used with PointNet (also has been in original publication)
- Also uses only symmetric aggregation function (like the max pooling in PointNet)
- Constructs local neighborhood graph using closed k-points in (feature) space
 (
 — Needs to define metric to measure distances in feature space)
- Applying convolution-like operations on the edges connecting neighboring pairs of points (called EdgeConvolution)
- The neighborhood graph is rebuild (in feature space) after each layer
 - \rightarrow Graph changes dynamically
 - \rightarrow No deterministic neighborhood relation
- Stacking this propagates local features over long distances and thus enables connection to global features



Our Network

- Use always the same k-neighbors
- Instead of finding neighbors in the feature space of each layer
- Made 'static' again by using always same neighborhood graph



Data Treatment

• Assume vertex to be known by other methods!

Use this to correct signals for time of flight

• Use special projection on unit-sphere:

Project PMT-position on unit sphere around vertex \rightarrow Angular position for each signal

Use time to modulate distance of point to origin

 \rightarrow Time deformed sphere around vertex

All signals surviving 2.75 ns time cut projected on unit sphere around vertex (red)



Influence of Training-Data



set	count	energy	position
training	210,000	3 MeV	center and z-axis *
validation	5,200	4 MeV	center
energy evaluation	8 x 5,200	1 - 8 MeV	center
vertex evaluation	5 x 5,200	4 MeV	z-axis: 0 to 10 meters

* 100,000 events at detector center, rest evenly distributed along z-axis

Final Results: JUNO

3 cm bias of vertex in flight direction is included (green): As expected from vertex reconstruction (that assumes only scintillation light)



- Not much directional information in JUNO
- But a change of strategy might help to improve (do not rely on vertex)

Results with better TTS

• Assuming 1 ns TTS for all PMTs (80% coverage)

 \rightarrow Theia-like instrumentation



Probably need to extract vertex and direction simultaneously to further improve this!

Results with better TTS

Assuming 1 ns TTS for all PMTs (80% coverage)

 \rightarrow Theia-like instrumentation

Gree Remark: Here only 3% Cherenkov-light (in WBLS much higher ratio)

- → Should be much easier in WBLS (work in progress!) (but will also depend on how fast the scintillation is in WBLS)
- The same DG-CNN could also be used to separate Cherenkov from scintillation signals



Blue

hift

rther

Goals at GeV Energies

- Non-ML methods: Full topological reconstruction can reveal many details
- But: Very computing intensive & lack robustness in some cases
- Question: Can ML do better?



Topological

How to do Something Similiar with ML?

Scenario used:

- Toy-MC simulating scintillation along random track with a high emission point (peak)
- No light attenuation or scattering (otherwise full LS model)
- Cubic detector with 4m edge length
- 100 PMTs with 1ns time resolution per wall (full coverage)

• Two output goals:

- Coordinates of start-, end- and peak-position
- Voxel reconstruction





ML Architecture For Shower Reconstruction

- First stage: Dynamic Graph CNN
- Second stage: Fully connected layers (standard CNN)



First Results: Shower (Peak) Finding



Next steps:

- Go to more realistic detector/simulation
- Look at more complicated events

Outlook: First Results Voxel Reconstruction

Using L1-regularization in loss function

Red: MC Truth Blue: Network output





-2.0 1 5 1.0 ...Z Workinprogressi -0.5 -1 0 1.0 0.5 -1.5 -1.0 -0.5 -0.5 0.0 X 0.5 1.0 -15 1.5 2.0 -2.0

Result of homogeneous network

Result after propagation layers

Result heterogeneous network (after training)

Event Classification with CNNs in ANNIE

- Water-Cherenkov detector (26 ton) operated in Booster-neutrino beam (at Fermilab)
- First neutrino experiment using LAPPDs
 - \rightarrow TTS ~ 0.1 ns, spatial resolution ~ 1mm
 - Assumed 24 LAPPDs in simulation for this study



Identifying CCQE-0 π Events

Two classification tasks:

Electron vs. muon

Single ring (SR) vs. multi ring (MR)



Result: Up to 5% increases of accuracy compared to classical methods \rightarrow Efficiency > 92% for each task (impurity <0.3%)

Summary/Conclusion

- Theia is a proposed detector
 - Large community interest, collaboration has been formed
 - White paper: M. Askins et al., Eur.Phys.J.C 80 (2020) 5, 416, arXiv:1911.03501
- ML learning is a central to reach its full potential
 - Particle Identification in pure LS already very successful
 - Profits a lot from fast timing, high granularity and large coverage
 - Separating Cherenkov-light will increase potential further
 - Directional reconstruction difficult in LAB+PPO+bis-MSB (MeV energies)

 \rightarrow The right cocktail (WBLS, slow LS, ...) will help a lot!

• **Goal:** Unlock power of C/S-ratio and ring-counting



Backup slides

Theia Interest Group



Most of these institutes joint the Theia proto-collaboration!

Bringing everything together



Community Interest

Site	Scale	Target	Measurements	Timescale
UChicago	bench top		fast photodetectors	Exists
CHIPS	10 kton	H2O	electronics, readout, mechanical infrastructure	2019
EGADS	200 ton		isotope loading, fast photodetectors	Exists
ANNIE	30 ton	H2O+Gd		Exists
WATCHMAN	l kton			2020
NuDot	l ton	LS	directionality	2018
Penn	30 L		light yield, timing, loading	Exists
SNO+	780 ton	(**0)23		2018
CHESS (LBNL)	bench top		signal separation, tracking, reconstruction / light yield, loading, attenuation	Exists
BNL	l ton	VVbLS		Exists

ANNIE CNN:Architecture



Influence of Directionality/Particle ID

- Cherenkov-light
 - Can reveal direction of solar neutrino events
 - And help particle identification
- Cutting events pointing away from sun reduces bkg
- Efficiency will strongly depend on scintillator & detector properties
 (MC-simulations for 50% coverage in water show 80% rejection of B-8 bkg with 75% signal efficiency)
 (R. Jiang & A. Elagin, arXiv:1902.06912: 65% coverage in LAB → >50% rejection with 70% signal efficiency)





Which events do we want to discriminate - and why?

α/β :

 α -background from natural radioactivity

p/β :

- supernova: e.g. elastic scattering channel with protons (signal) vs β^- decay from ⁸⁵Kr, ²¹⁰Bi and ¹⁴C
- elastic scattering channel with protons vs elastic scattering channel with electrons
- **geo-neutrino**: suppress background from ${}^{13}C(\alpha,n){}^{16}O$ interactions, which can result in elastic scattering on proton

e^{+}/e^{-} :

- **IBD**: measure ⁸He and ⁹Li yield in interactions with cosmic muons (both β^{-})
- solar: MSW-transition region (2 MeV 4 MeV) partly dominated by β^+ decays from ¹⁰C -

e^{-}/γ :

solar: external gammas from outside of the CD, high loss of exposure due to deep fiducial volume cut

Decay Schemes I-130 & Cs-136



Cherenkov-Light Separation by Wavelength

ntum Efficiency

ã

• Using dichroic filter

(transmitting above or below a certain threshold, reflecting the rest)

• Optimal Cut for LAB-PPO (2g/I): 450 nm Full description in T. Kaptanoglu et al., JINST 14 (2019) no.05, T050





Long-Baseline Physics with Theia

- Ring-imaging for long-baseline physics
- SK & HK improved reconstruction methods a lot
 → Theia competitive long-baseline



Solar Neutrinos with Theia

10

10

10

10

Events / 0.05 MeV

- Directionality very potent tool
- Also powerful: Discrimination pointlike & non-point-like events (like C-10)
- Li-loading can make CC-channels accessible $_{^7Li + \nu_e \rightarrow \ ^7Be + e^-}$ (Q = 862 keV)



Theia White Paper, to be published soon (Courtesy to R. Bonventre &G.D. Orebi Gann)

CC 7Be

ES 7Be

---- ES CNO

CC CNO

-Sum

-CC 8B

--- ES 8B

— CC pep

--- ES pep

Supernova Neutrinos in Theia

- Core-collapse SN at 10kpc
- Opens new physics window:
 - Test SN models
 - Information about MH
 - Multi-messenger astronomy
 - Early warning with precise pointing (< 1°) $\frac{(\text{NCO})^{-16}O(\nu,\nu)^{-16}O^*}{(\nu,\nu)^{-16}O^*}$

Huge statistics + Flavour information

Reactio	on	Rate
(IBD)	$\bar{\nu}_e + p \to n + e^+$	19,800
(ES)	$\nu + e \rightarrow e + \nu$	960
$(\nu_e O)$	${ m ^{16}O}(u_e,e^-){ m ^{16}F}$	340
$(\bar{\nu}_e O)$	${}^{16}{ m O}(\bar{\nu}_e,e^+){}^{16}{ m N}$	440



1.100

DSNB with Theia

Combines Neutrino signal of past SN

Encoded information from:

- Star formation rate
- Average core-collapse neutrino spectrum

Advantage Theia:

Pulse-shape discrimination, Ring-Counting C/S-ratio



Theia White Paper

DSNB with Theia

Combines Neutrino signal of past SN

Encoded information from:

- Star formation rate
- Average core-collapse neutrino spectrum
- Advantage Theia:
 - Pulse-shape discrimination, Ring-Counting C/S-ratio



17kt fiducial mass

Nucleon Decay with Theia

- Triple coincidence: $p \rightarrow \overline{v} K^+ \rightarrow Kaon decay \rightarrow decay of decay product$
- Invisible decay of oxygen nucleus:

 $n \rightarrow 3\nu \rightarrow$ One 6.18 MeV γ from excited nucleus



Theia White Paper

Geo-Neutrinos with Theia

- Thousands of Geo-neutrino events per year
 - \rightarrow Precise measurement of Th & U components in spectrum
- Expected rate would be 2s greater than the KamLAND rate after 1 year (at SURF)
 - \rightarrow First evidence for surface variation of flux possible



$0\nu\beta\beta$ in Theia: Expected Endpoint Spectra



• **Resulting Sensitivity** (90% C.L.):

$$\begin{split} \mathbf{Te}: \ T_{1/2}^{0\nu\beta\beta} > 1.5\times 10^{28} \ \mathrm{y}, \ m_{\beta\beta} < 5.4 \ \mathrm{meV} \\ \mathbf{Xe}: \ T_{1/2}^{0\nu\beta\beta} > 2.7\times 10^{28} \ \mathrm{y}, \ m_{\beta\beta} < 4.8 \ \mathrm{meV} \end{split}$$

After 10 years

(Signal loss due to B-8 rejection not included yet)

Topological Reconstruction at High Energies

- Can make dE/dx and complex event structure visible (even in pure LS)
- Needs fast timing & good time resolution



In unsegmented large-volume liquid scintillator detectors



See B.W. et al., arXiv:1803.08802

My Basic Idea

Assumption:

- One known reference-point (in space & time)
- Almost straight tracks
- Particle has speed of light
- Single hit times available

Concept:

• Take this point as reference for all signal times

The Drop-like Shape

Signal time = particle tof + photon tof



Working Principle Part I Summary

- For each signal:
 - Time defines drop-like surface
 - Gets smeared with time profile (scintillation & PMT-timing)
 - Weighted due to spatial constraints (acceptance, optical properties, light concentrator, ...)
- \rightarrow Spatial p.d.f. for photon emission points







Working Principle Part II



Loss Function

• Loss-function: TF.Losses.Cosine_Distance

(Tensor-standart-Loss-Function)

• Minimizes: $L = \sqrt{(\cos(\theta - 1))^2}$

 θ = angle between reconstructed and true direction

- True direction:
 - Used start direction of event
 - Not average flight direction of electrons

Remark: Resolution for the later could be better, but at least for background reduction we are interested in the former. For correction of energy reconstruction it would be better to take average direction.

Improving Liquid Properties

Development of scintillating liquids

- WBLS (Brookhaven NL, JGU Mainz, TU Munich)
- Isotope loading (BNL, MIT) (Li,B,Ca,Zr,In,Te,Xe,Pb,Nd,Sm,Ge,Yb)
- Oil-diluted LS (JGU Mainz)
- **Characterization** (Brookhaven NL, JGU Mainz, TU Munich, ...)
 - Optical properties (Emission, attenuation, ..)
 - Timing properties (Time spectrum, ortho-positronium, ...)
- Filtering methods (Attenuation, radiopurity)
 - Nanofiltration (UC Davis)



New WBLS: JHU Mainz & TU Munich







Finding the Right Cocktail



light separation in LAB-PPO

CHErenkov Scintillation Separation

10/07/20

LAPPD propagation effects

PRC 95 055801 (2017)

CHESS

Photo Sensor Development (Fast & Efficient & Affordable & High Granularity)

• LAPPDs (Fast timing & high granularity)





Commercially available now (Incom Inc.)^{time (psec)} Used in ANNIE + R&D at U Chicago

• HQE 20" PMTs (Efficient & affordable)







Modular PMTs (Good compromise of everything)



Water-Cherenkov Test Beam Experiment

 SiPM + active light guide (Very efficient + increasing affordability)



Cherenkov-Light Separation by Wavelength

