Applying Machine Learning for $\beta\beta$-decay Identification

Andrey Elagin
University of Chicago

Work in progress

In collaboration with
E.Toropov (Carnegie Mellon), I.Vukotic (Chicago),
S.Fraker (MIT), L.Winslow (MIT)
C.Grant (Boston), A.Li (Boston)

THEIA Meeting, UC Davis, April 13, 2018
Outline

• Brief reminder on previous work: the Spherical Harmonics Analysis
• Current status of applying machine learning techniques to $0\nu\beta\beta$-decay identification

In this talk I focus specifically on $0\nu\beta\beta$-decay THEIA physics case
Double-Beta Decay Kinematics

- Lots of “back-to-back” (large angle) events
- Most of electrons are above Cherenkov threshold
Is it possible to use Cherenkov light for 0νββ-decay reconstruction in a liquid scintillator detector?

This question was formulated in: JINST 7 (2012) P07010; PRD87 (2013) 071301

[In chronological order, comments on omissions and latest publications are welcome]

Yes, we can use Cherenkov light in 0νββ-decay events, but this requires
- fast photo-detectors (red sensitive photo-cathode helps) [see JINST 9 (2014) 06012]
- slow scintillators [see NIMA 849 (2017) 102]
Cherenkov Light Comes First

Scintillation emission is slower
- Longer wavelengths travel faster
- Cherenkov light arrives earlier

- $370 \text{ nm} \Rightarrow 0.191 \text{ m/ns}$
- $600 \text{ nm} \Rightarrow 0.203 \text{ m/ns}$

$\sim 2 \text{ ns}$ difference over 6.5m distance
Using Directionality of Early Light

**Simulation:**
- single electrons along X-axis at the center of 6.5m sphere
- KamLAND scintillator

**Reconstruction:**
- WCSim adapted for low energy

Directionality “survives” some detector effects
Vertex resolution is promising

Directionality is a handle on $^8$B events
Directionality or Topology?

Idealized event displays: no multiple scattering of electrons, all PEs, QE=30%

0νββ-decay

Cherenkov PEs

S power spectrum

Scintillation PEs

Spherical harmonics analysis

\[ f(\theta, \varphi) = \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{\ell} f_{\ell m} Y_{\ell m}(\theta, \varphi). \]

Rotation invariant power spectrum

\[ S_{ff}(\ell) = \sum_{m=-\ell}^{\ell} |f_{\ell m}|^2 \]
Early Light Topology

"Realistic" event displays: early PEs only, KamLAND PMTs QE: Che~12%, Sci~23%

Early PE: 0νββ-decay

Cherenkov PEs

Early PE: 8B event

Scintillation PEs

Spherical harmonics analysis

Rotation invariant power spectrum

\[ f(\theta, \varphi) = \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{\ell} f_{\ell m} Y_{\ell m}(\theta, \varphi). \]

\[ S_{ff}(\ell) = \sum_{m=-\ell}^{\ell} |f_{\ell m}|^2 \]

S power spectrum

Normalized power

Multipole moment
Multipole moment $l=0$

Simulation details:

- 6.5m radius detector
- Scintillator model from KamLAND simulation (1 ns rise time)
- TTS=100 ps, 100% area coverage, QE(che) $\sim 12$, QE(sci) $\sim 23$
- Central events only

Key parameters determining separation of $0\nu\beta\beta$-decay from $^8$B:

- Scintillator properties (narrow spectrum, slow rise time)
- Photo-detector properties (fast, large-area, high QE, red-sensitive)
Current implementation of the spherical harmonics analysis requires vertex reconstruction. This limits applicability to multi-vertex events such as $^{10}$C and gammas.
$0\nu\beta\beta$ vs $^8B$

Vertex res 5cm, events within $R<3m$
Scintillation rise time 1 ns

Vertex res 5cm, events within $R<3m$
Scintillation rise time 5 ns

First quantitative demonstration of benefits of slow scintillators for $0\nu\beta\beta$-decay event topology reconstruction

For details see NIM A849 (2017) 102

Background rejection factor = 2
@ 70% signal efficiency

Background rejection factor = 3
@ 70% signal efficiency
Can We Use Machine Learning?

In the early light, the difference between one-track and two-track events often can be seen by eye. Computer should be able to learn to see the difference.

There are plenty of tools to try. A 2D Convolutional Neural Network seems to be a natural choice to start.

- Feed “photos” of theta-phi plane into a convolutional neural network.
- Bin the photos in time and treat those timing bins as color channels.
“Photos” of $0\nu\beta\beta$ and $^8\text{B}$ Events

- Two-track event
  - "color" channel $\Delta T_1$
  - "color" channel $\Delta T_2$
  - "color" channel $\Delta T_3$

- One-track event
  - "color" channel $\Delta T_1$
  - "color" channel $\Delta T_2$
  - "color" channel $\Delta T_3$
ConvNet vs Spherical Harmonics

$0\nu\beta\beta$-decay vs $^8B$: central events only

While currently there is only small improvement over spherical harmonics, ConvNet does not use any information about vertex -> important simplification in dealing with off-center events as well as with gammas, positrons, and $^{10}C$ backgrounds

Work in progress
ConvNet for $0\nu\beta\beta$-decay vs $\gamma$

$0\nu\beta\beta$-decay vs $\gamma$: any events with $R<3m$

Work in progress

Classifier output

No explicit vertex reconstruction

AUC = 84%
Why it works better on $\gamma$?

More differences in overall timing distributions

$\gamma$ Compton scatters $\rightarrow$ may have multiple vertices

PE arrival times relative to the very first PE

Also $\sim50\%$ of $e^+$ from $^{10}\text{C}$ will form an ortho-positronium $\rightarrow$ even longer delay
Next Steps

- Try current ConvNet on $^{10}$C events

- Check correlation between spherical harmonics and ConvNet classifier outputs
  - Likely to be highly correlated given similar performance, but this needs to be checked
  - If there is independent information use ConvNet for vertexing and feed it into spherical harmonics, then add spherical harmonics to the ConvNet classifier

- Try more “color” channels
  - Currently using 0.5 ns timing bins, smaller bins are computationally more challenging

- Try more sophisticated ML techniques
  - Consulting with experts in computer vision

- Apply to more realistic detector settings
  - SNO+ and/or KamLAND-Zen
  - THEIA detector parameter optimization (including mirrors)

- Are we using all information?
  - Wavelength
  - Polarization
  - Photon angular information (the Distributed Imaging method: PRD 97 052006)
Conclusions

• A rather simple implementation of ML techniques looks promising

• Current performance on $0\nu\beta\beta$-decay vs $^8$B is similar to spherical harmonics analysis

• Key advantage of ML methods is that they do not explicitly depend on vertex reconstruction
Back-up
My e-mail exchange with Jenni Kotila:

"...The angular correlation is basically the $a^{(1)}/a^{(0)}$, where $a^{(i)}$ are defined in Eq. (24) for $2nbb$ and in Eq. (51) for $0nbb$. In case of $0nbb$ only thing that matters are the electron wavefunctions but in case of $2nbb$ there are these additional factors that are a combination of $\langle K_N \rangle$ and $\langle L_N \rangle$, that are defined in Eq. (23) and include the electron energies, the neutrino energies and the closure energy. So even with small neutrino energies, for example $e_1=0.749Q$, $e_2=0.249Q$, $w_1=0.002Q$, $w_2=0$ a factor of 0.4329 is obtained. Regarding the question about the situation for different isotopes, the closure energy entering the equations is different for each isotope and can be approximated by $1.12A^{(1/2)}$ MeV..."
Directionality of Early Photons

Cherenkov photons from center of 6.5m-radius sphere: TTS=100 psec

Cosine of angle between the photoelectron hit and the original electron direction after the 34 ns cut. Both Cherenkov and scintillation light are included. Note the peak at the Cherenkov angle.

C.Aberle, A.Elgin, H.Frisch, M.Wetstein, L.Winslow
2014 JINST 9 P06012
What About Lower Energies?

Light yield: Cherenkov vs scintillation

\[ \frac{1}{2} Q (^{116}\text{Cd}) = 1.4 \text{ MeV} \]

\[ \frac{1}{2} Q (^{48}\text{Ca}) = 2.1 \text{ MeV} \]
$\beta\beta$ vs $^{8}\text{B}$

Vertex res 5cm, events within $R<3\text{m}$
Sci rise time 1 ns

$I_{\text{overlap}} = 0.79$
$^{0\nu}\beta\beta$ vs $^{8}\text{B}$

Vertex res 5 cm, events within $R<3\text{m}$
Sci rise time 5 ns

$I_{\text{overlap}} = 0.64$
Off-Center Events

\[ x' = \frac{ax}{\sqrt{1 + a_x^2 + a_y^2}} \cdot R \]

\[ y' = \frac{a_y}{\sqrt{1 + a_x^2 + a_y^2}} \cdot R \]

\[ z' = \frac{a_z}{\sqrt{1 + a_x^2 + a_y^2}} \cdot R \]

\[ ax = x_{hit} - x_{v+\mathbf{x}} \]

\[ ay = y_{hit} - y_{v+\mathbf{x}} \]

\[ az = z_{hit} - z_{v+\mathbf{x}} \]
**0νββ-decay vs 10C**

*two-track vs a “complicated” topology*

---

**10C decay chain:**

- $^{10}C$ final state consist of a positron and gamma
  (e$^+$ also gives 2x0.511MeV gammas after losing energy to scintillation)
- Positron has lower kinetic energy than $0νββ$ electrons
- Positron scintillates over shorter distance from primary vertex
- Gammas can travel far from the primary vertex

$^{10}C$ background can be large at a shallow detector depth

---

Diagram by Jon Ouellet

---

**10C vs 0νββ-decay: photons arrival time profile**

TTS=100 ps
**0νββ-decay vs \(^{10}\text{C}\)**

Time profile for events uniformly distributed within the fiducial volume, R<3m. Vertex resolution of 3cm is assumed.

**Photons count in early light sample**

Spherical harmonics help here too.

**Disclaimer:** there are other handles on \(^{10}\text{C}\) that are already in use (e.g., muon tag, secondary vertices). Actual improvement in separation power may vary.