European Research Council
Established by the European Commission
Supporting top researchers from anywhere in the world

VNIVERSITAT DE VALÈNCIA

IFIC INSTITUT DE FÍSICA CORPUSCULAR

Deep Learning and

The NEXT Experiment

IML Machine Learning Working Group
CERN
May 24, 2016

Josh Renner
IFIC/University of Valencia, Spain

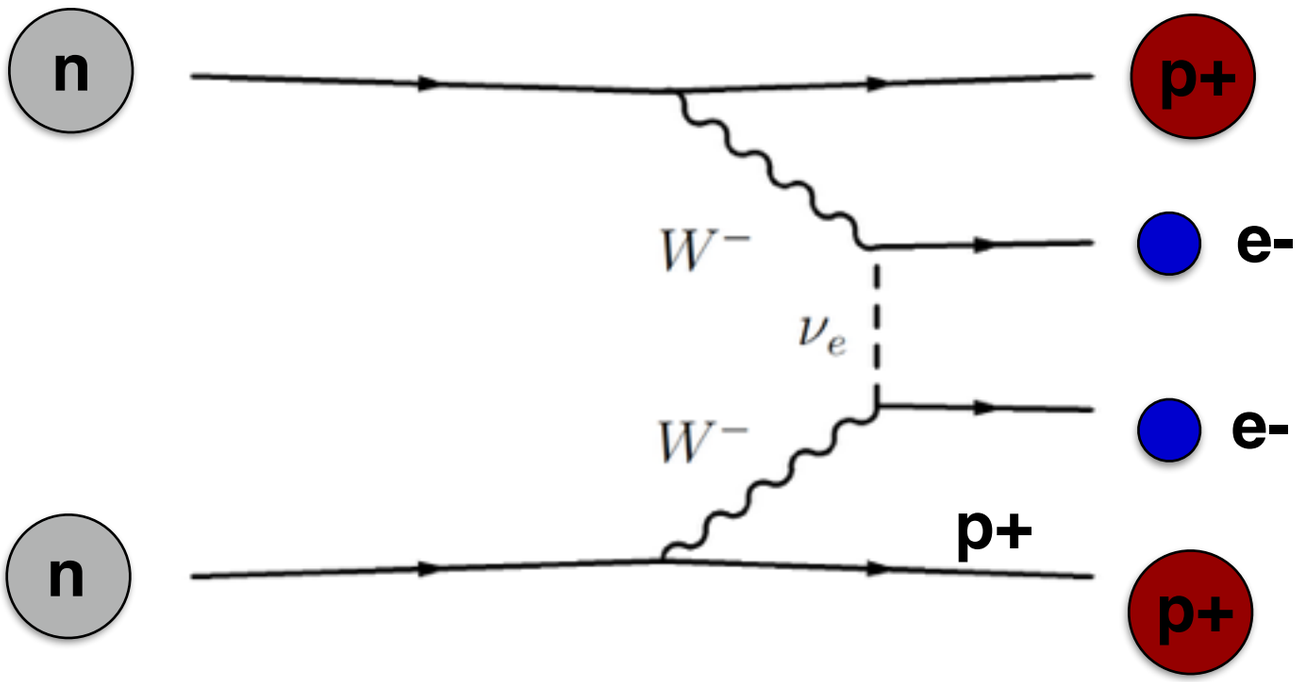
Neutrino **E**xperiment with a **X**enon **T**PC

LSC
Laboratorio Subterráneo de Canfranc

NEXT at Canfranc

- search for neutrinoless double-beta decay (0vßß)
- final proposed phase: 100 kg Xe, enriched to ^{136}Xe (90%)
- high pressure gas, electroluminescent TPC

Why 0vßß?



$^{136}Xe \longrightarrow {}^{136}Ba^{++} + 2e^-$

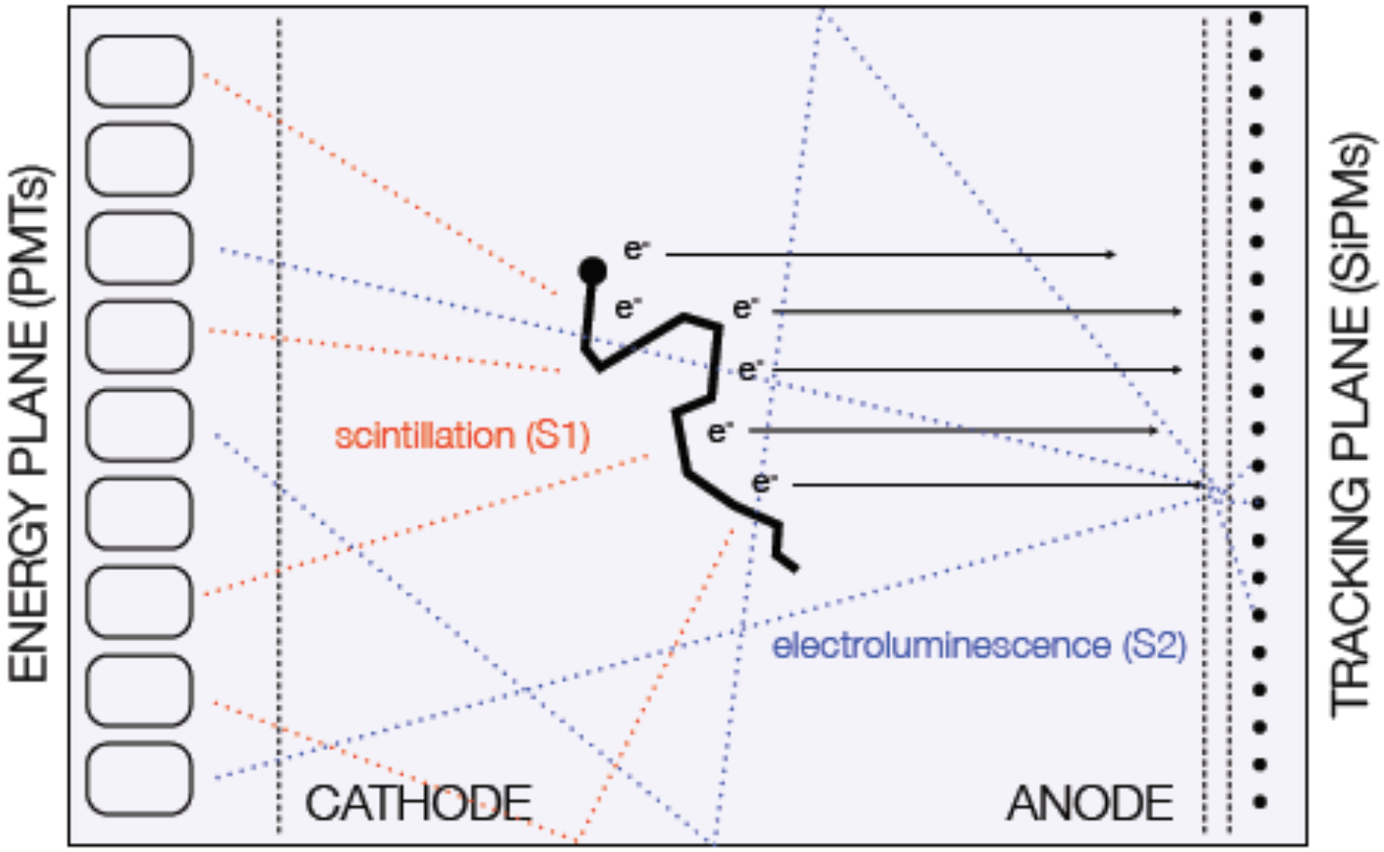
- 2 simultaneous ß decays
- 2 anti-neutrinos emitted become a virtual neutrino

An observation of 0vßß would:

- demonstrate that neutrinos are Majorana
  - new mass scale / mass generation mechanism?
- imply total lepton number violation
  - explanation of matter/anti-matter asymmetry (leptogenesis)?
- provide information on neutrino hierarchy and mass scale

An Electroluminescent HPXe TPC

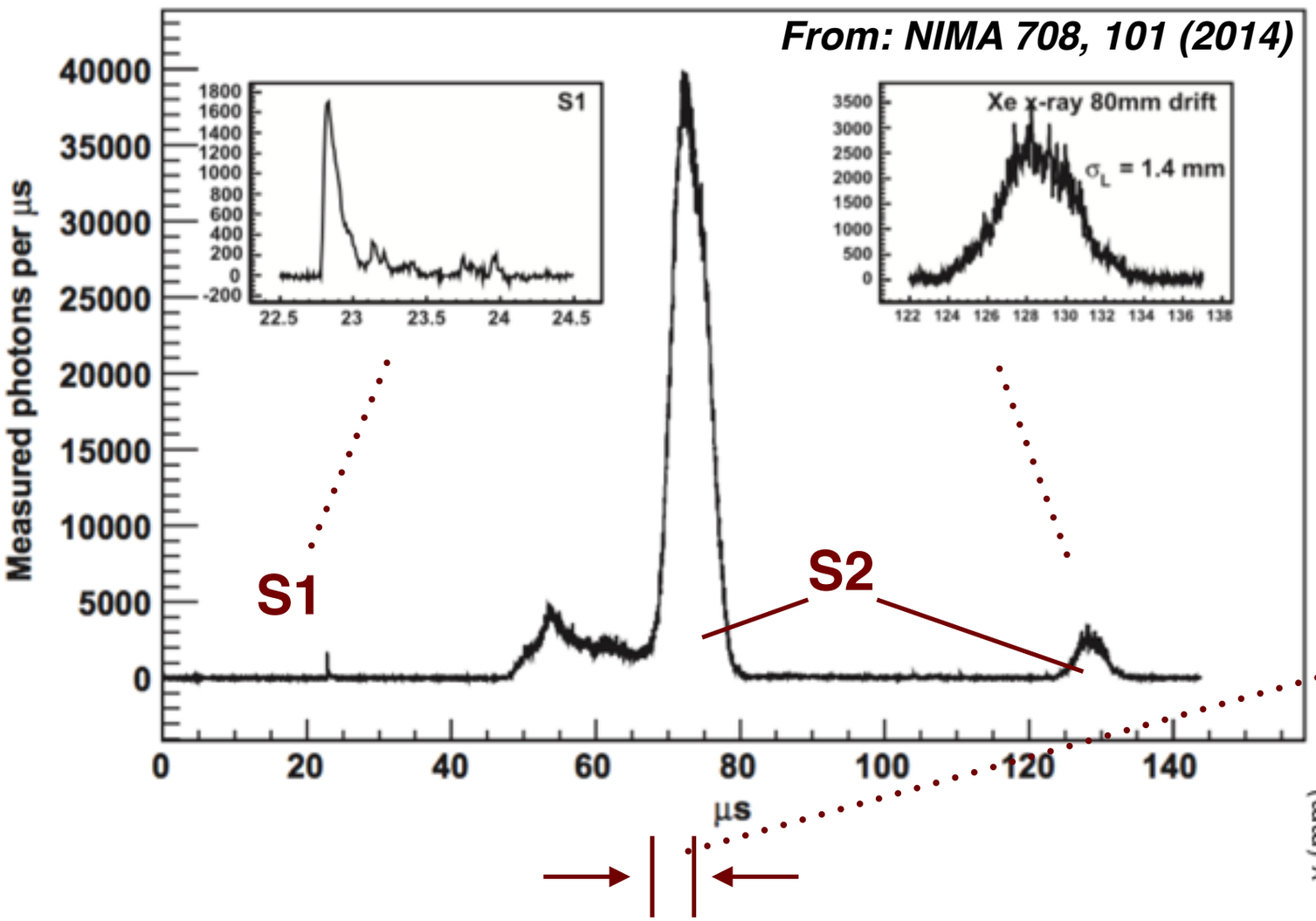
- ionization track drifted to electroluminescent readout plane
- S2 light detected by silicon photomultipliers (SiPMs) in the *tracking* plane, and PMTs in the *energy* plane
- initial S1 light measured by PMTs to provide t_0



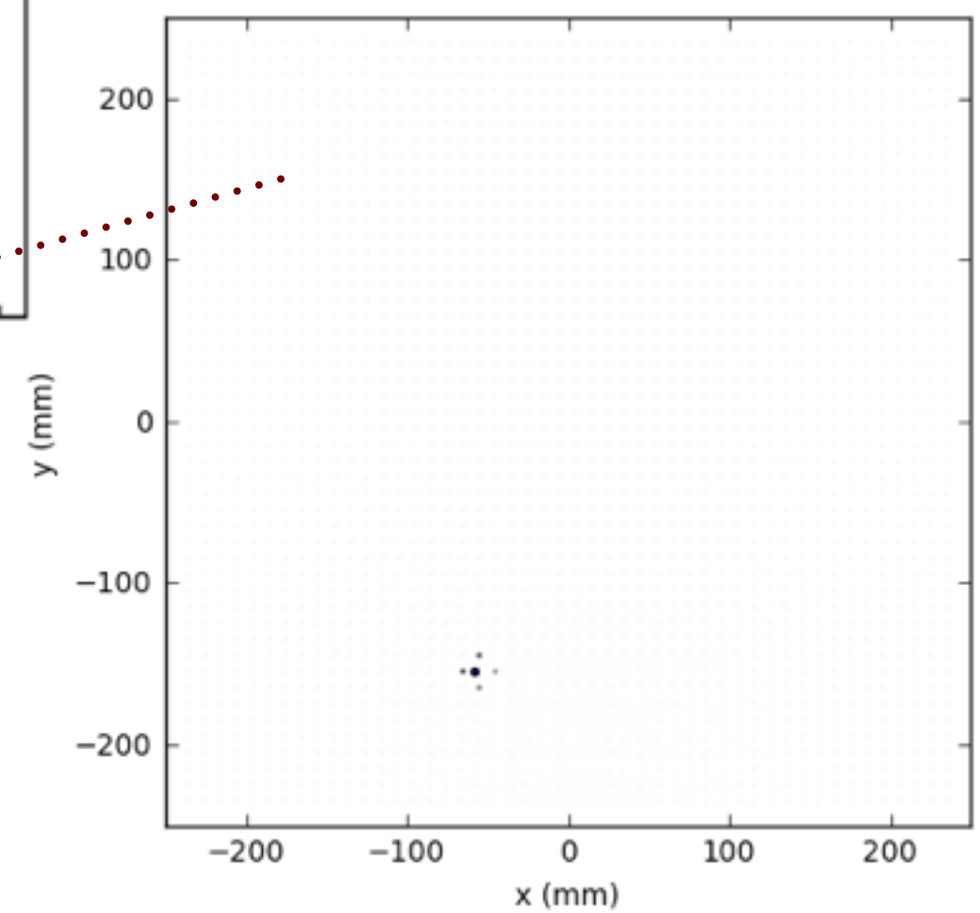
- **Simultaneous precise energy measurement and 3D event reconstruction**

Signals observed in an EL-TPC

Observed by the *energy plane* (PMTs)



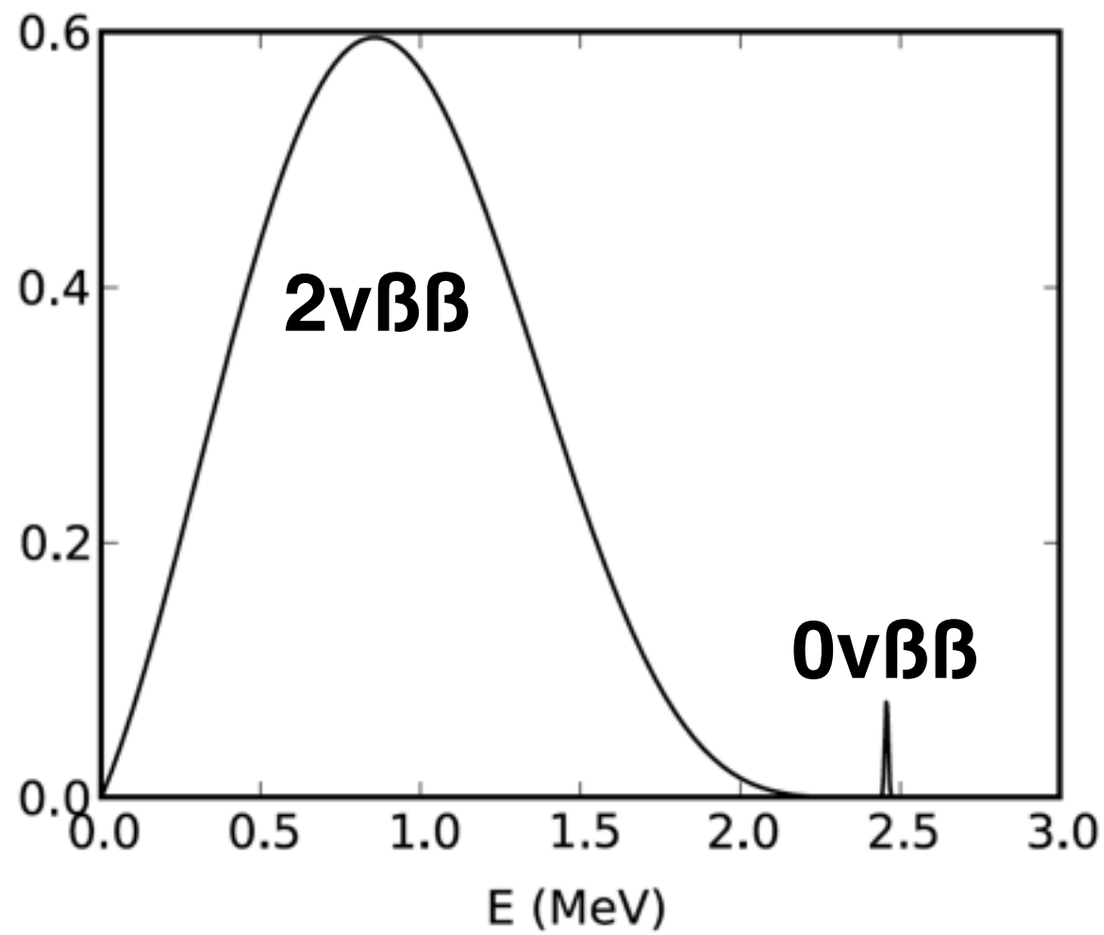
Observed by the *tracking plane* (SiPMs)



- S2 sliced into time (z) bins and SiPM measured signal can be used to determine (x,y)

What are we looking for?

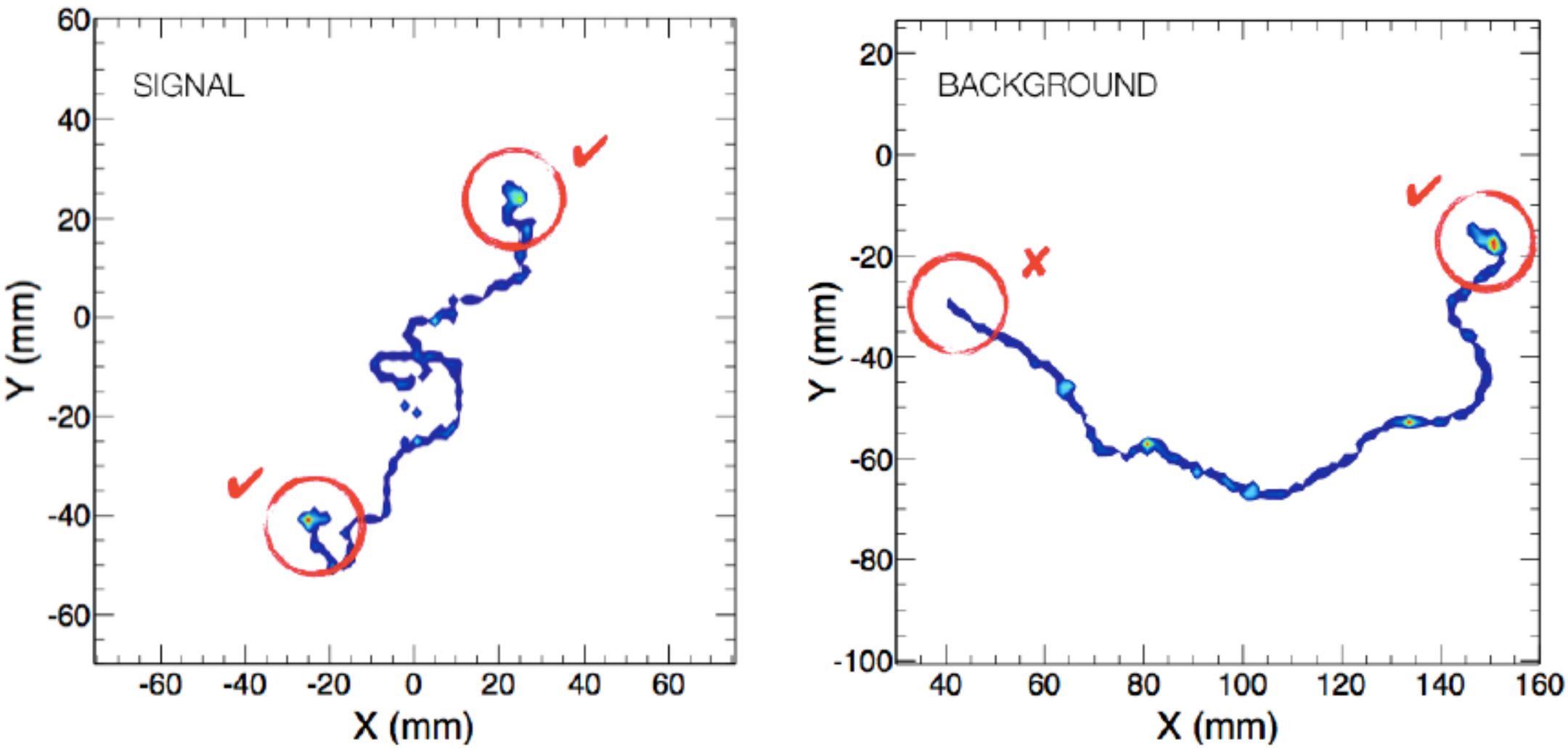
0vßß would yield 2 electrons of total energy = $Q_{ßß}$



- But we could still get background events that are not 0vßß but still fall into the energy peak

Topological signature in NEXT

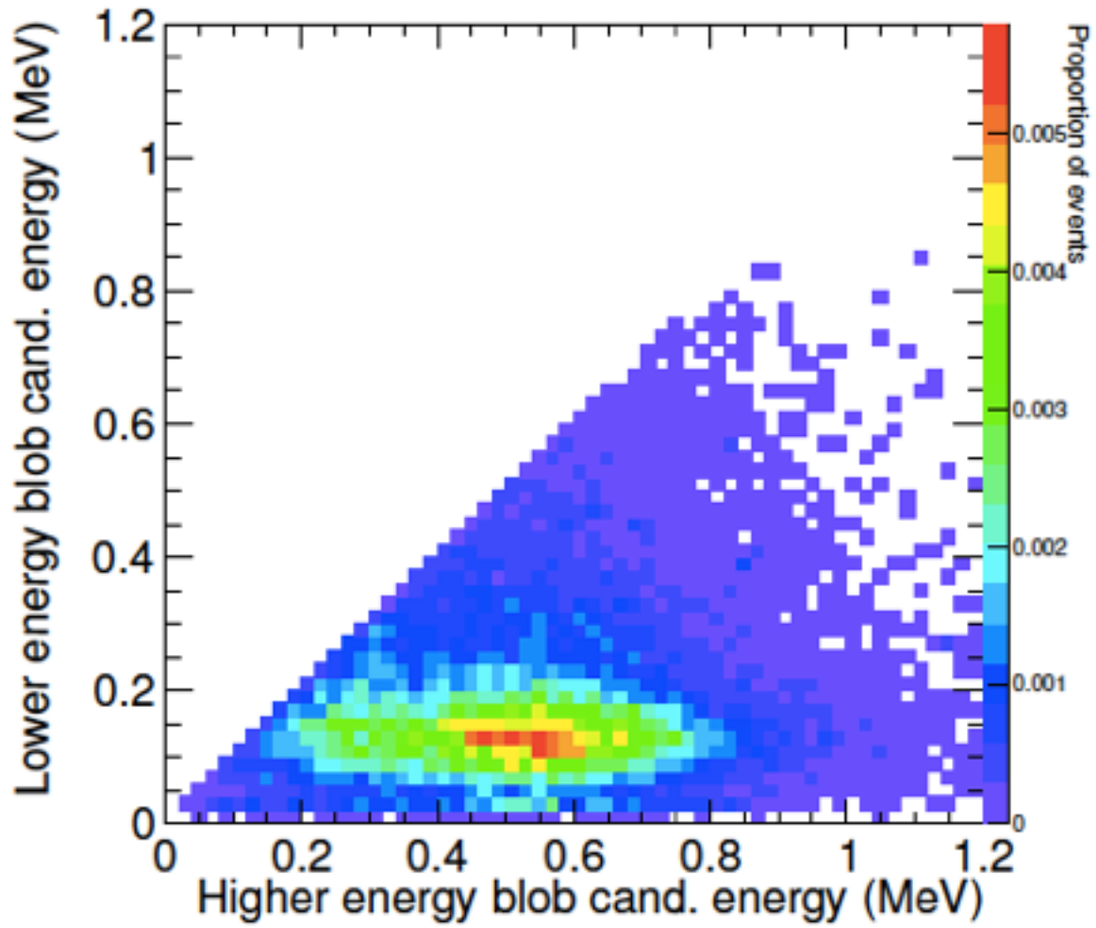
JHEP 01, 104 (2016) [arXiv:1507.05902]

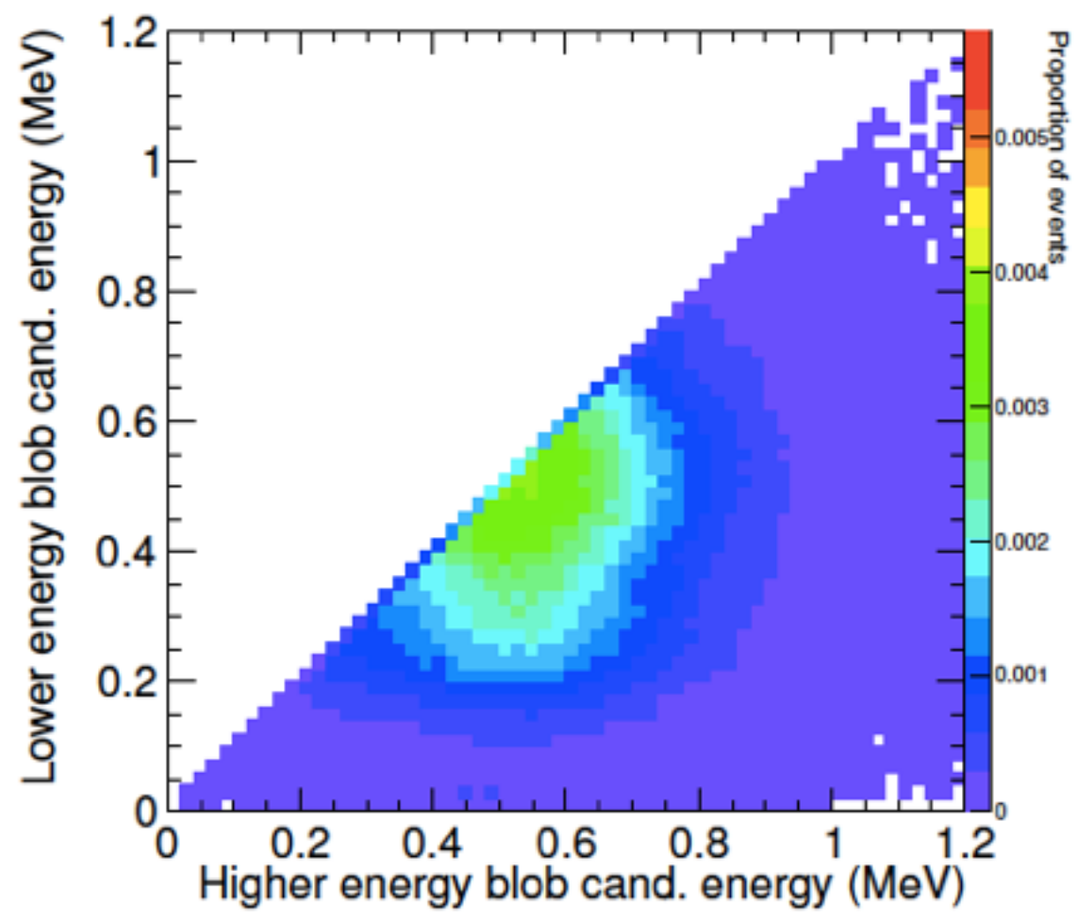


- Energetic electron leaves a high-density deposition at the end of its track (Bragg peak)
- Results in distinct topological signatures for **signal** and **background** events of the same energy

Topological signature in NEXT

JHEP 01, 104 (2016) [arXiv:1507.05902]





- Sum of energies in region enclosing ends of track provides main discriminant
- Geant4 Monte Carlo: (left) 2.44 MeV gammas from ^{214}Bi decay, (right) 0νββ events

NEXT-100: Background Rejection

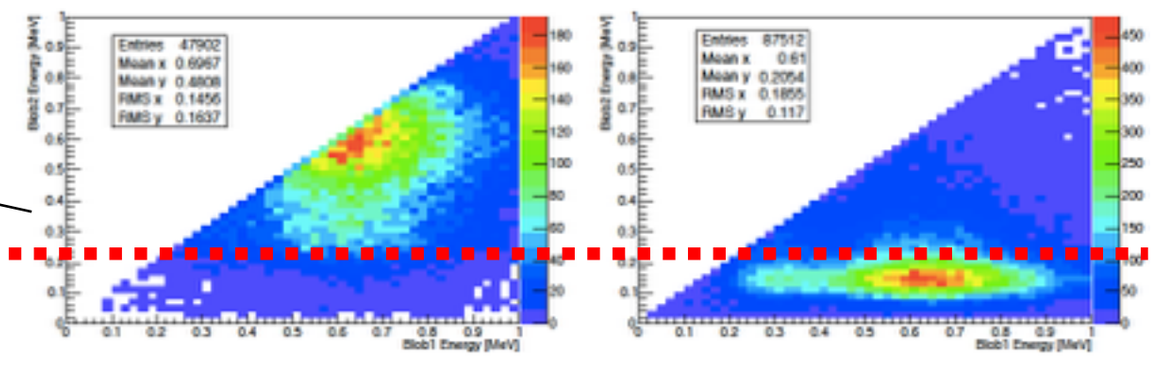
- **Monte Carlo-based study:** arXiv:1609.06202
  - Full NEXT-100 geometry, 10^5 initial signal events in active region, BG events (10^9 initial ^{208}Tl and 10^{10} initial ^{214}Bi) launched from field cage

	Signal Events		BG Events (^{208}Tl)		BG Events (^{214}Bi)	
Cut	$2\times2\times2$	$10\times10\times5$	$2\times2\times2$	$10\times10\times5$	$2\times2\times2$	$10\times10\times5$
(Initial events)	1.0	1.0	1.0	1.0	1.0	1.0
Energy	7.59×10^{-1}	7.59×10^{-1}	2.27×10^{-3}	2.27×10^{-3}	1.42×10^{-4}	1.42×10^{-4}
Fiducial	6.71×10^{-1}	6.68×10^{-1}	1.19×10^{-3}	1.17×10^{-3}	8.62×10^{-5}	8.54×10^{-5}
Single-Track	3.75×10^{-1}	4.79×10^{-1}	7.90×10^{-6}	1.81×10^{-5}	3.84×10^{-6}	8.75×10^{-6}
Classification*	3.23×10^{-1}	3.67×10^{-1}	7.70×10^{-7}	2.41×10^{-6}	2.90×10^{-7}	9.59×10^{-7}

Reconstruction granularity (mm^3)

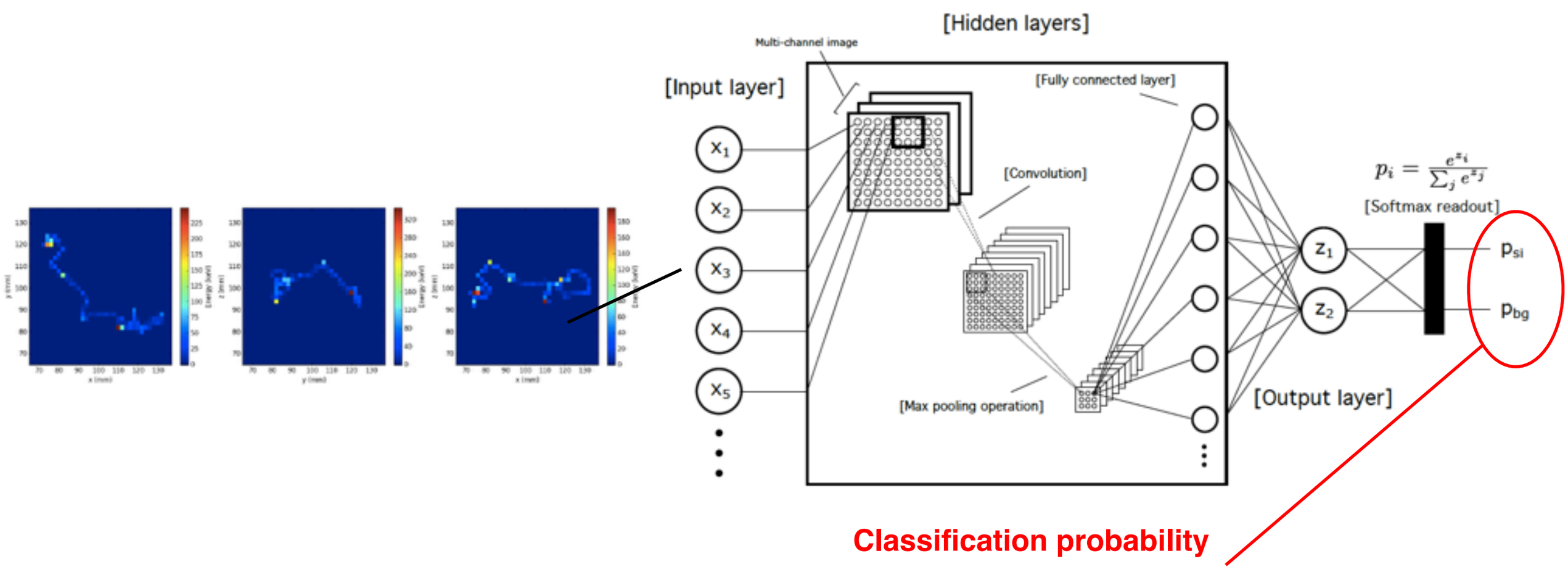
- 2 x 2 x 2 — optimistic, “low diffusion”
- 10 x 10 x 5 — current realistic case

- **Energy cut:** select events in E in [2.4,2.5] MeV
- **Fiducial cut:** < 10 keV deposited within 2 cm of walls
- **Single-track cut:** events with 1 connected track
- **Classification:** “blob”-based analysis



NEXT-100: Background Rejection

- **Replace classification cut with a deep neural network (DNN)**
  - Trained on several hundred thousand independently generated events
  - Events converted to 2D color images (R, G, B) = (xy, yz, xz) projections and used to train GoogLeNet on DIGITS (Caffe interface)

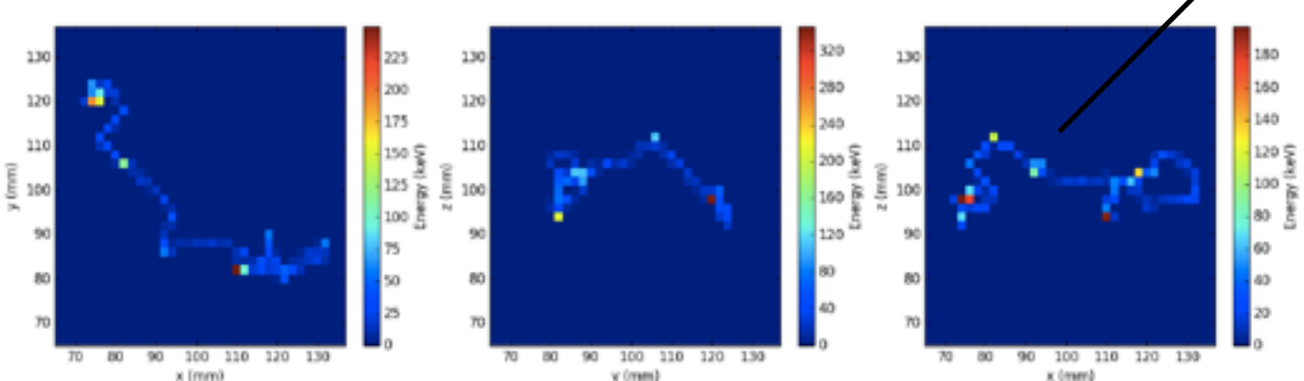


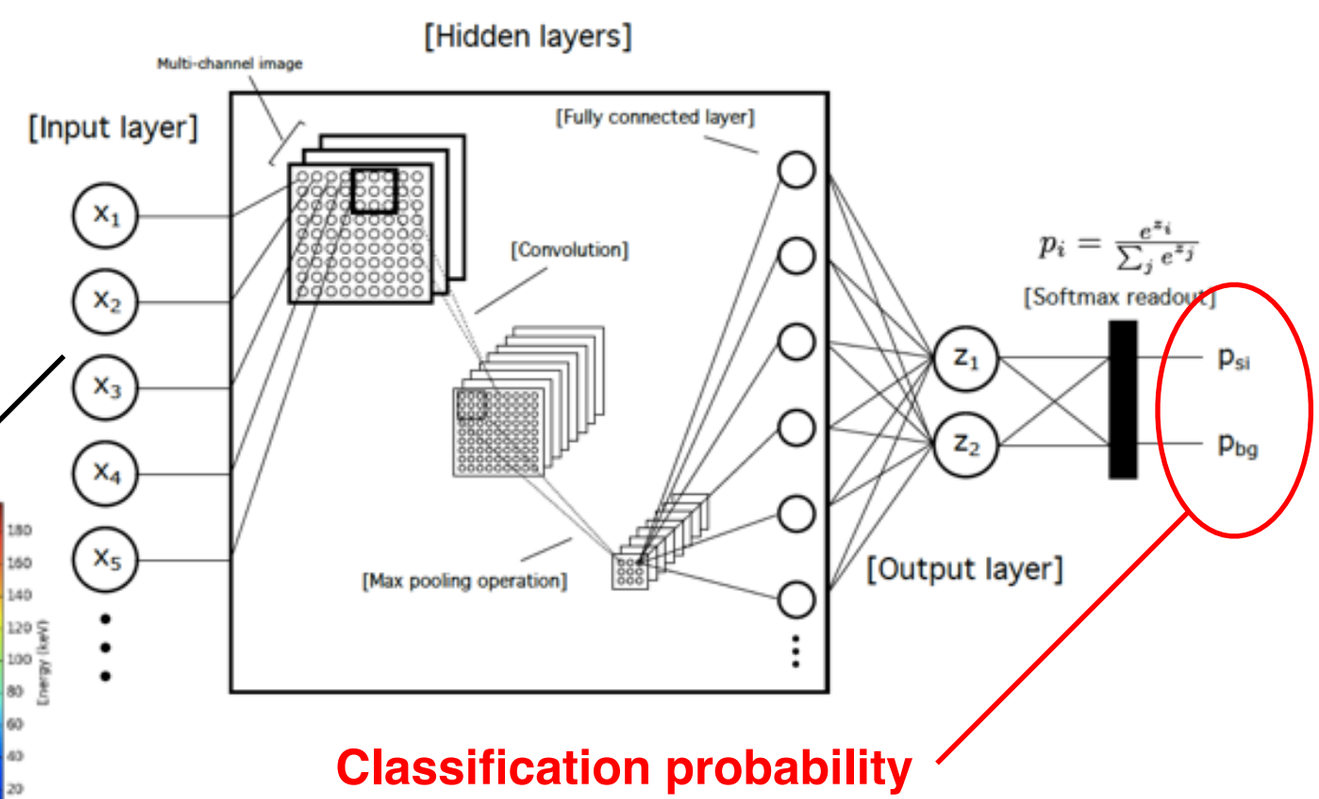
NEXT-100: Background Rejection

- **Monte Carlo-based study:** arXiv:1609.06202

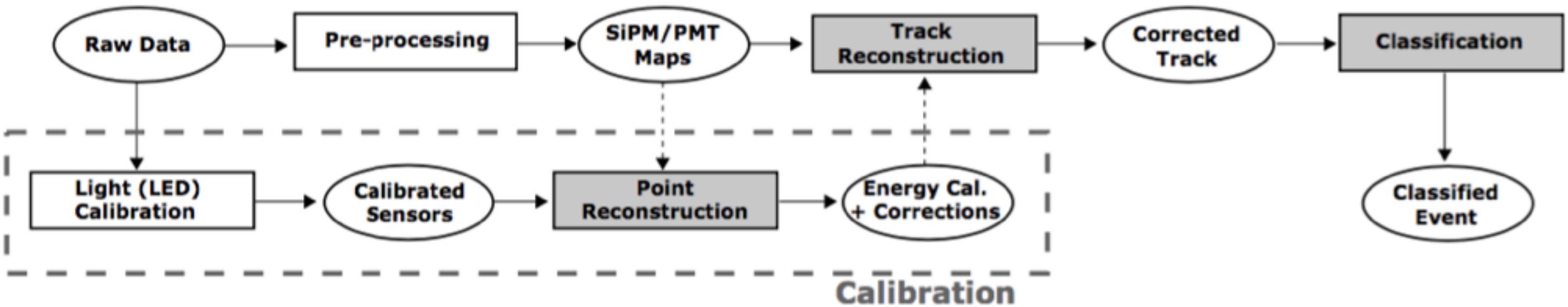
Cut	Signal Events		BG Events (^{208}Tl)		BG Events (^{214}Bi)	
	$2\times2\times2$	$10\times10\times5$	$2\times2\times2$	$10\times10\times5$	$2\times2\times2$	$10\times10\times5$
(Initial events)	1.0	1.0	1.0	1.0	1.0	1.0
Energy	7.59×10^{-1}	7.59×10^{-1}	2.27×10^{-3}	2.27×10^{-3}	1.42×10^{-4}	1.42×10^{-4}
Fiducial	6.71×10^{-1}	6.68×10^{-1}	1.19×10^{-3}	1.17×10^{-3}	8.62×10^{-5}	8.54×10^{-5}
Single-Track	3.75×10^{-1}	4.79×10^{-1}	7.90×10^{-6}	1.81×10^{-5}	3.84×10^{-6}	8.75×10^{-6}
Classification*	3.23×10^{-1}	3.67×10^{-1}	7.70×10^{-7}	2.41×10^{-6}	2.90×10^{-7}	9.59×10^{-7}
Classification (DNN)	**3.23×10^{-1}**	**3.67×10^{-1}**			**1.80×10^{-7}**	**8.22×10^{-7}**

- **Replace classification cut with a deep neural network (DNN)**
  - Trained on several hundred thousand independently generated events
  - Events converted to 2D color images (R, G, B) = (xy, yz, xz) projections





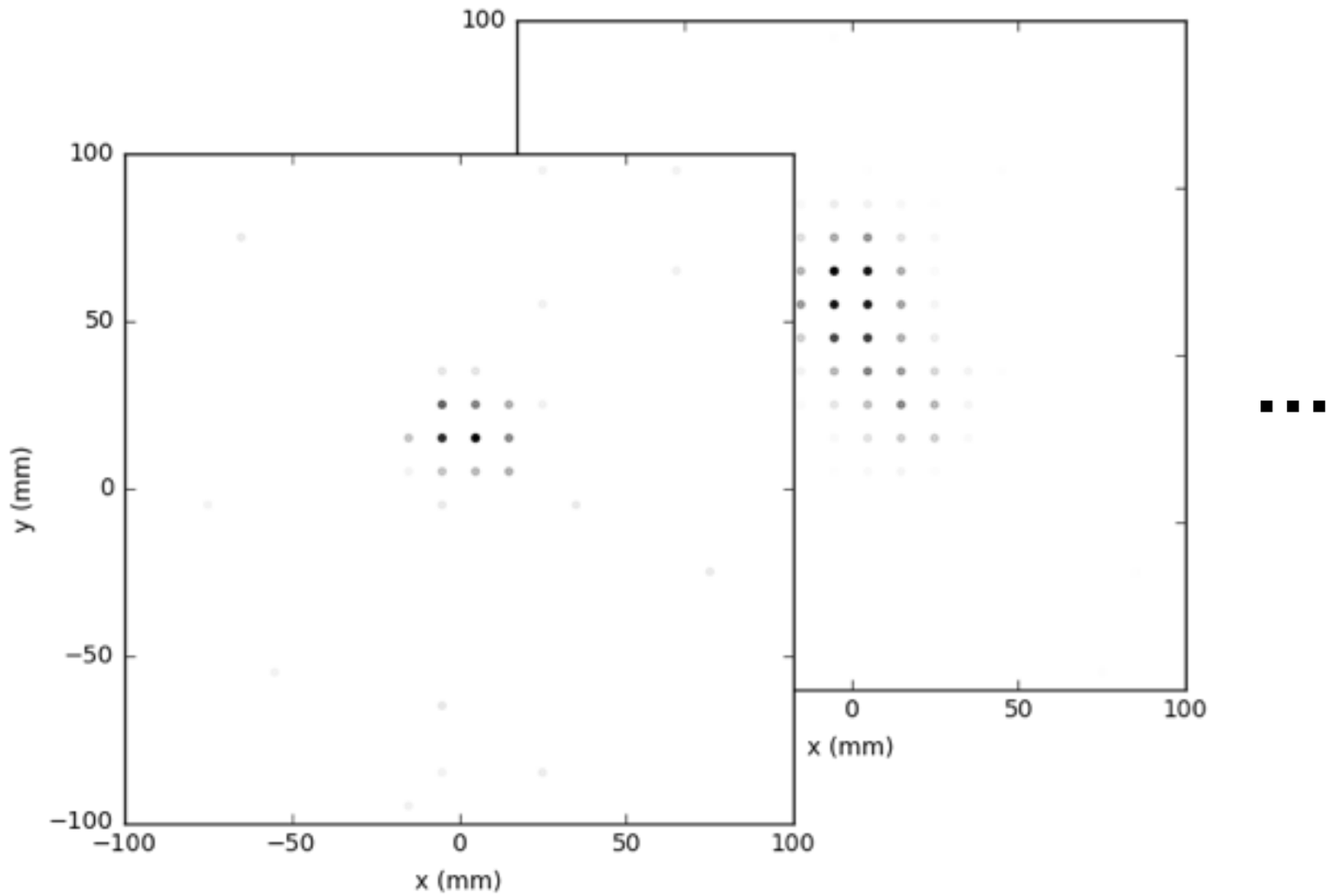
Examples (preliminary): Further applications of deep learning in NEXT



- **Point reconstruction:** take point-like events of a known energy and compute their x-y position. This along with their measured energy gives a calibration to correct for geometrical (x,y) effects in light collection.
- **Track reconstruction:** produce an image of the entire track (this was previously only considered using "classical" algorithms)
- **Classification:** signal vs. background as already discussed.

*** Note:** the following experiments were done using **Keras (Tensorflow)**

Classification: why reconstruct first?



- Use a **3D convolutional net** (several 3D convolutional layers + one 128-neuron fully connected layer + one 1-neuron readout layer) where the input is 20x20x60 matrix of SiPM signals
- See: https://jerenner.github.io/next-dnn-topology

Classification: why reconstruct first?

- **Potential improvement in a "full" (sensors simulated) Monte Carlo, but requires further study**

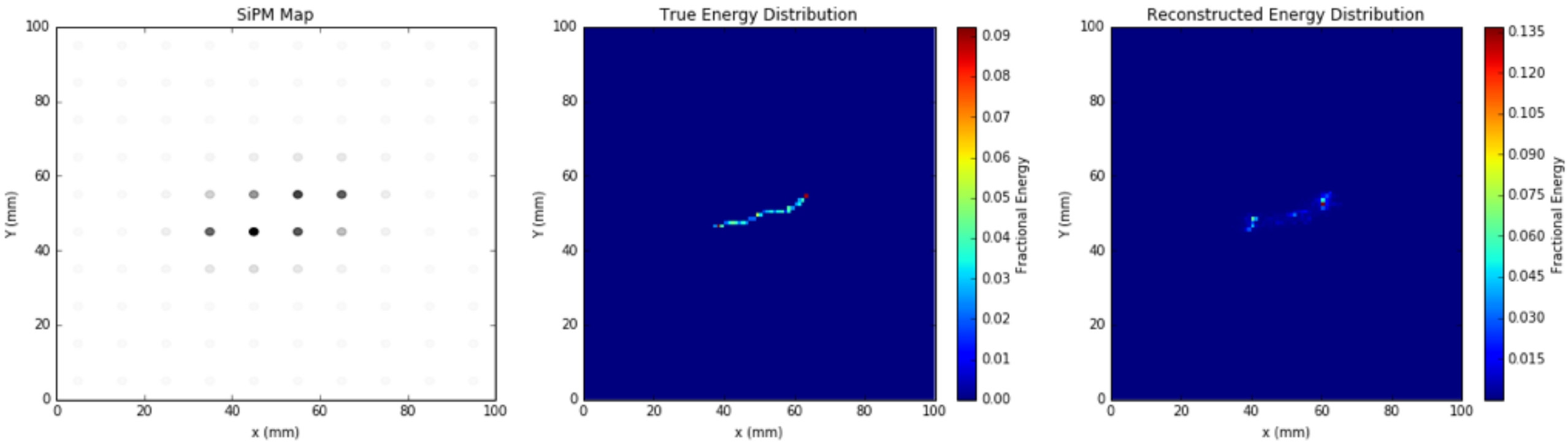
- **Classical analysis**

	Single-electron		0vbb	
Input Events	49997	100,00 %	49997	100,00 %
Fiducial	35780	71,56 %	36088	72,18 %
Single-track	14437	28,88 %	16406	32,81 %
Blobs	1869	3,74 %	11251	22,50 %

- **DNN analysis**

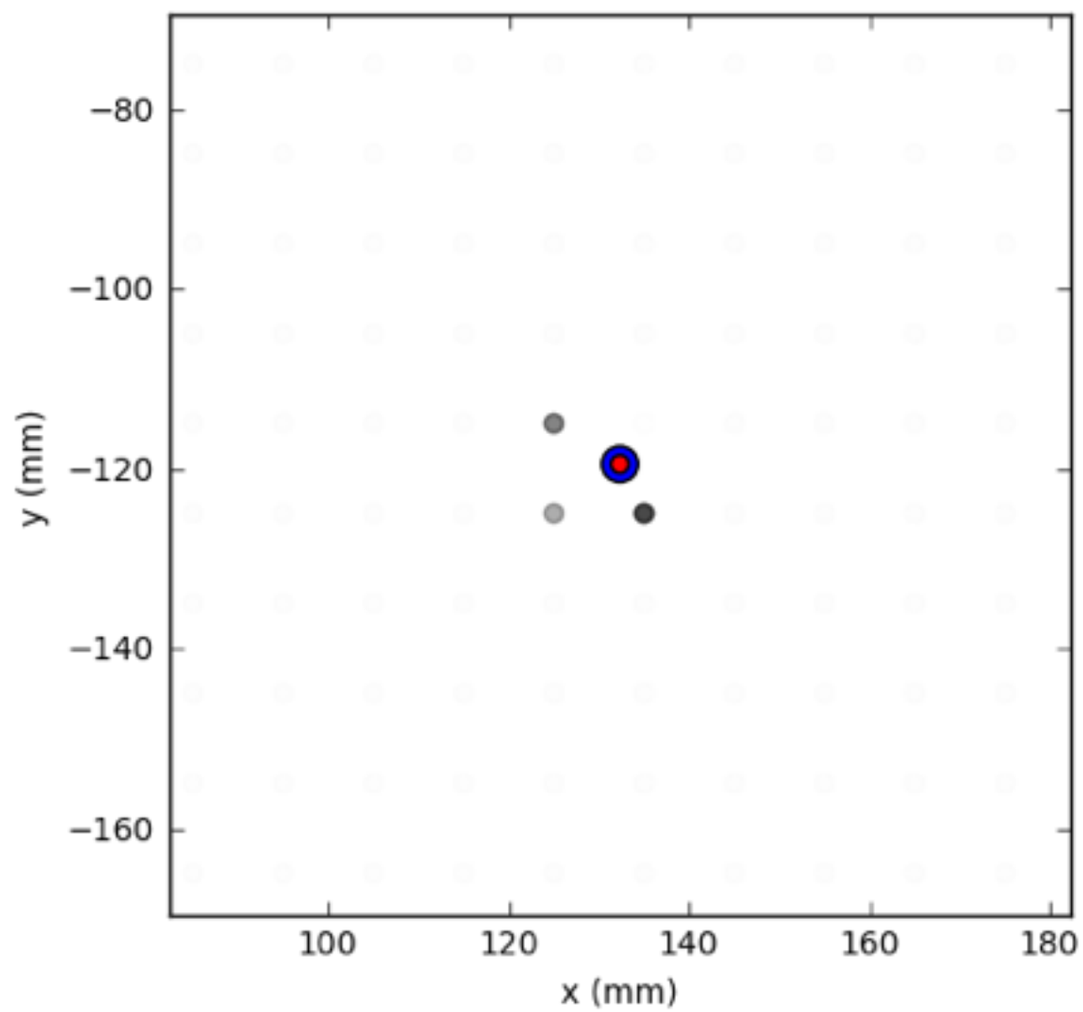
	Single-electron		0vbb	
Input Events	49997	100,00 %	49997	100,00 %
Fiducial (window)	44657	89,32 %	47412	94,83 %
DNN	1100	2,20 %	11251	22,50 %

Track Reconstruction



- Simulate SiPM response for many events and construct the true energy distribution (1 mm resolution) for z-slices of some size
- Train a 2D CNN to reconstruct the distribution from the simulated pattern of SiPM signals
- Probably only useful to produce an image of the track (classification, which is the ultimate goal, would ideally skip this step)

Point Reconstruction



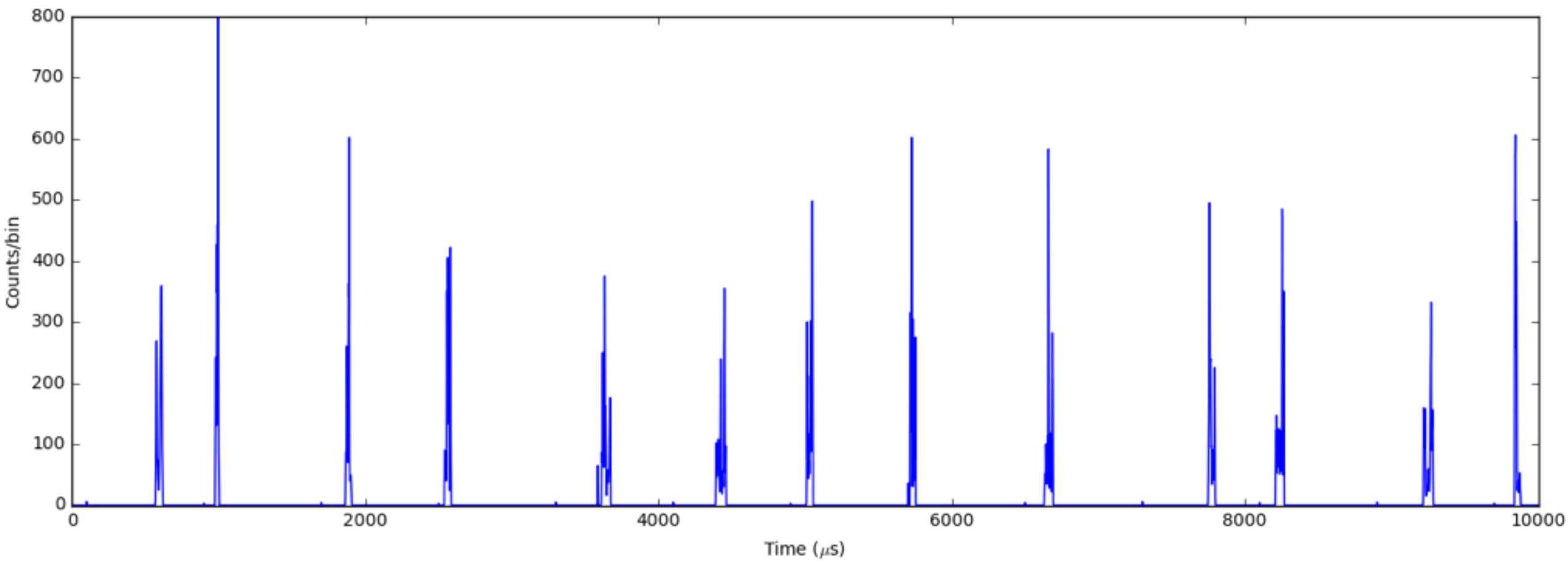
- Simulate point-like (40 keV) events and resulting single 2D projection of SiPM signals
- Reconstruct the (x,y) coordinates using a multi-layer fully-connected NN (above example: blue is true point, red is reconstructed)
- Later use the NN to reconstruct (x,y) for 40 keV electron tracks produced by ^{83m}Kr decay in actual detector

Summary

- Deep learning can be applied to several aspects of the analysis in NEXT and has shown potential to yield improvement
- Several MC-based studies have been carried out, but tests involving actual data are required
- The intermediate-scale (10kg) TPC called NEXT-NEW is currently running in the LSC (Canfranc); we expect to have initial data to which these ideas can be applied in the near future

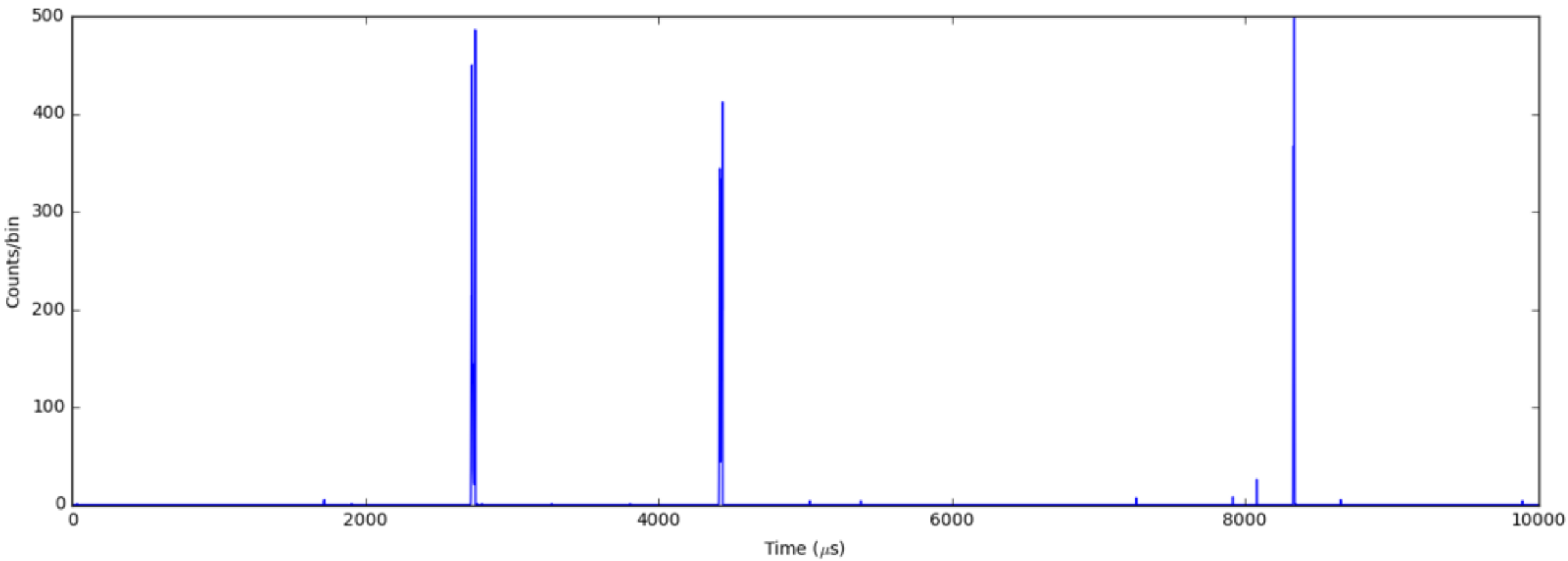
“Deep Analysis”

- Extract features from raw signals with deep learning
- For example, consider a concatenation of 1000 Monte Carlo events (sample of this shown below):



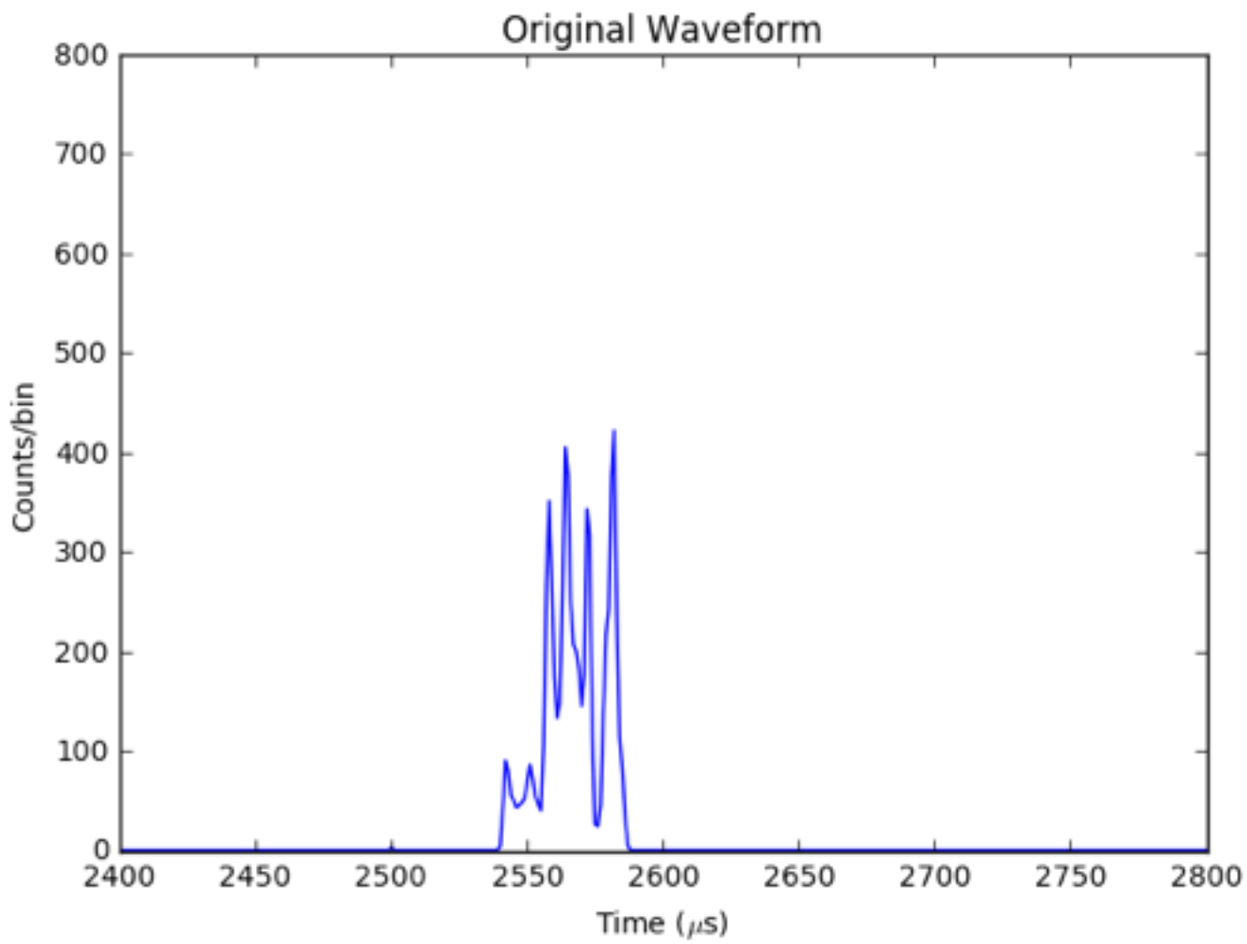
“Deep Analysis”

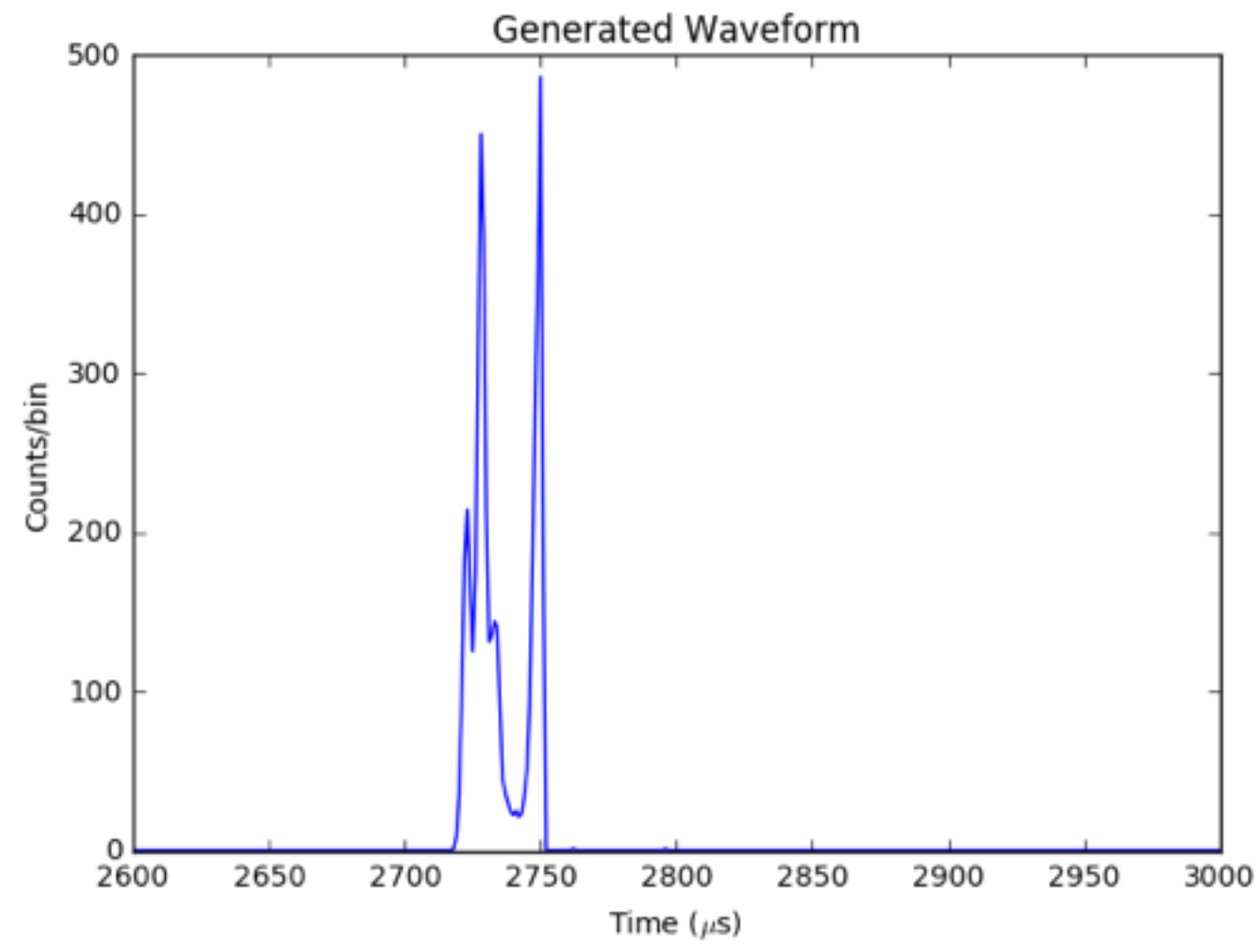
- Use a LSTM (long short-term memory) recurrent neural network (RNN) to predict the next value given sequence of 10 values
- Can generate a waveform with similar features (though still some discrepancies)



“Deep Analysis”

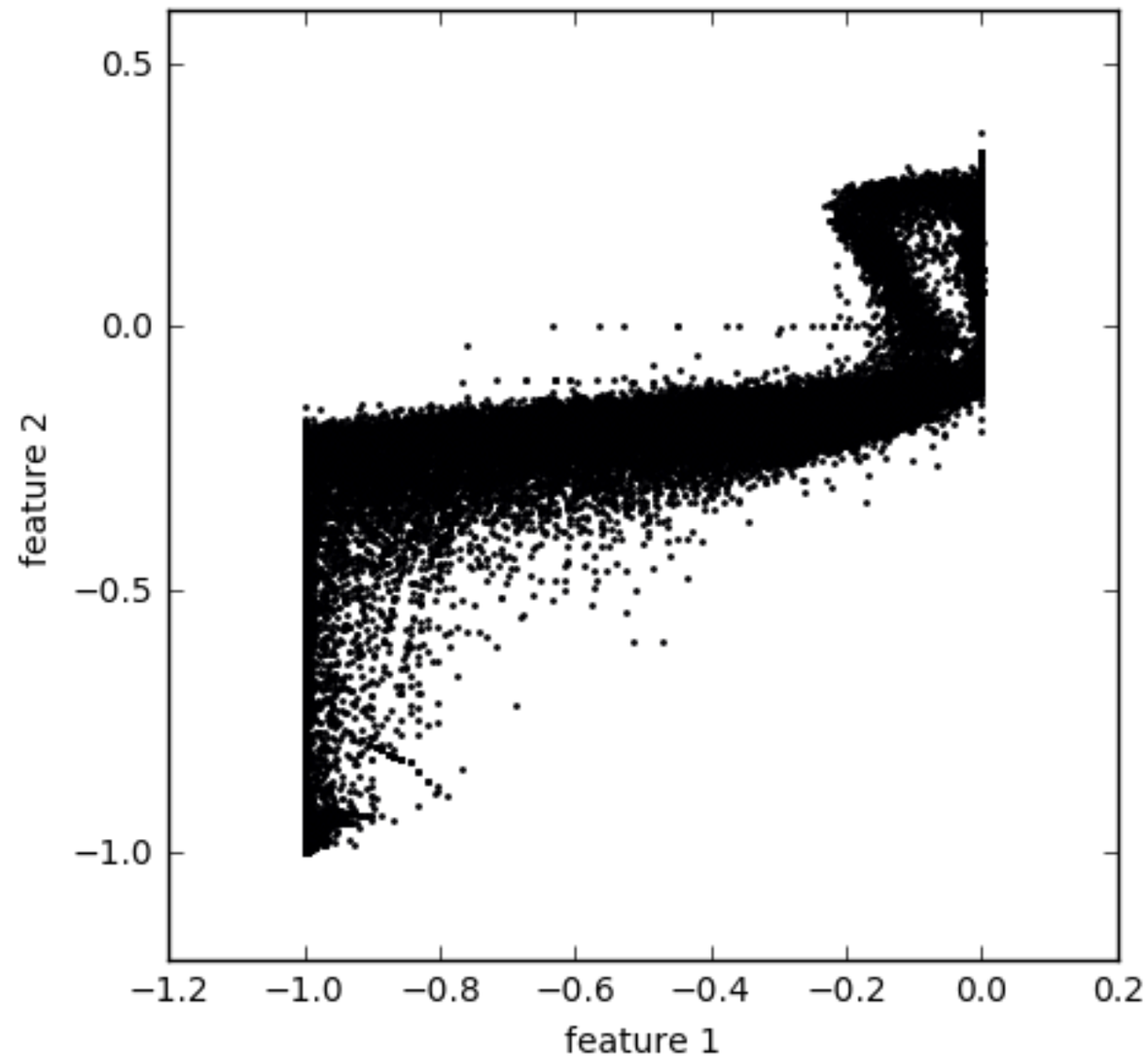
- Use a LSTM (long short-term memory) recurrent neural network (RNN) to predict the next value given sequence of 10 values
- Can generate a waveform with similar features (though still some discrepancies)





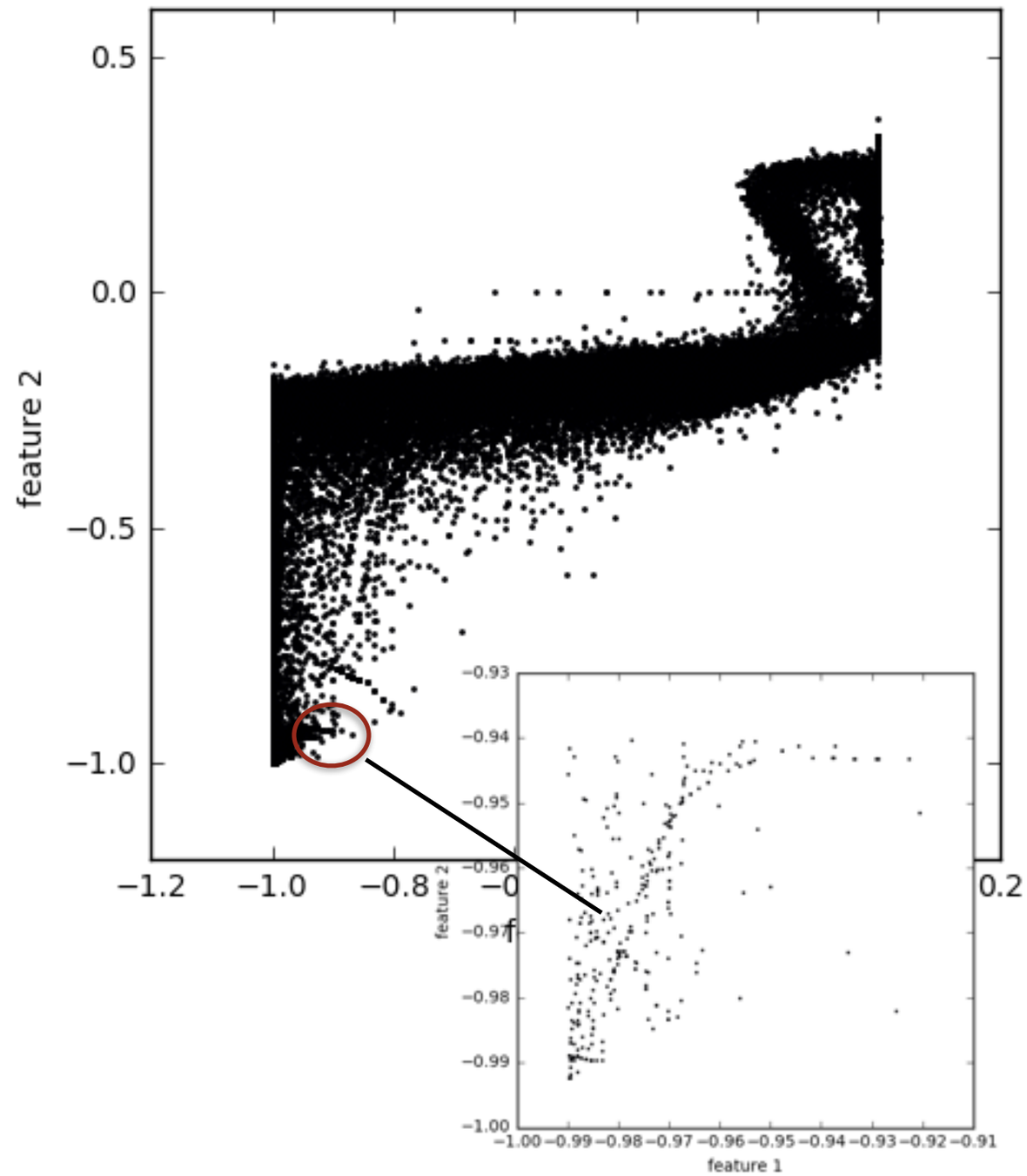
“Deep Analysis”

- Is this even useful?
- Note that if we have a machine that can do this, it must have learned something about the system. How can we extract this information?
  - Try to extract it from the outputs of specific neurons in the network

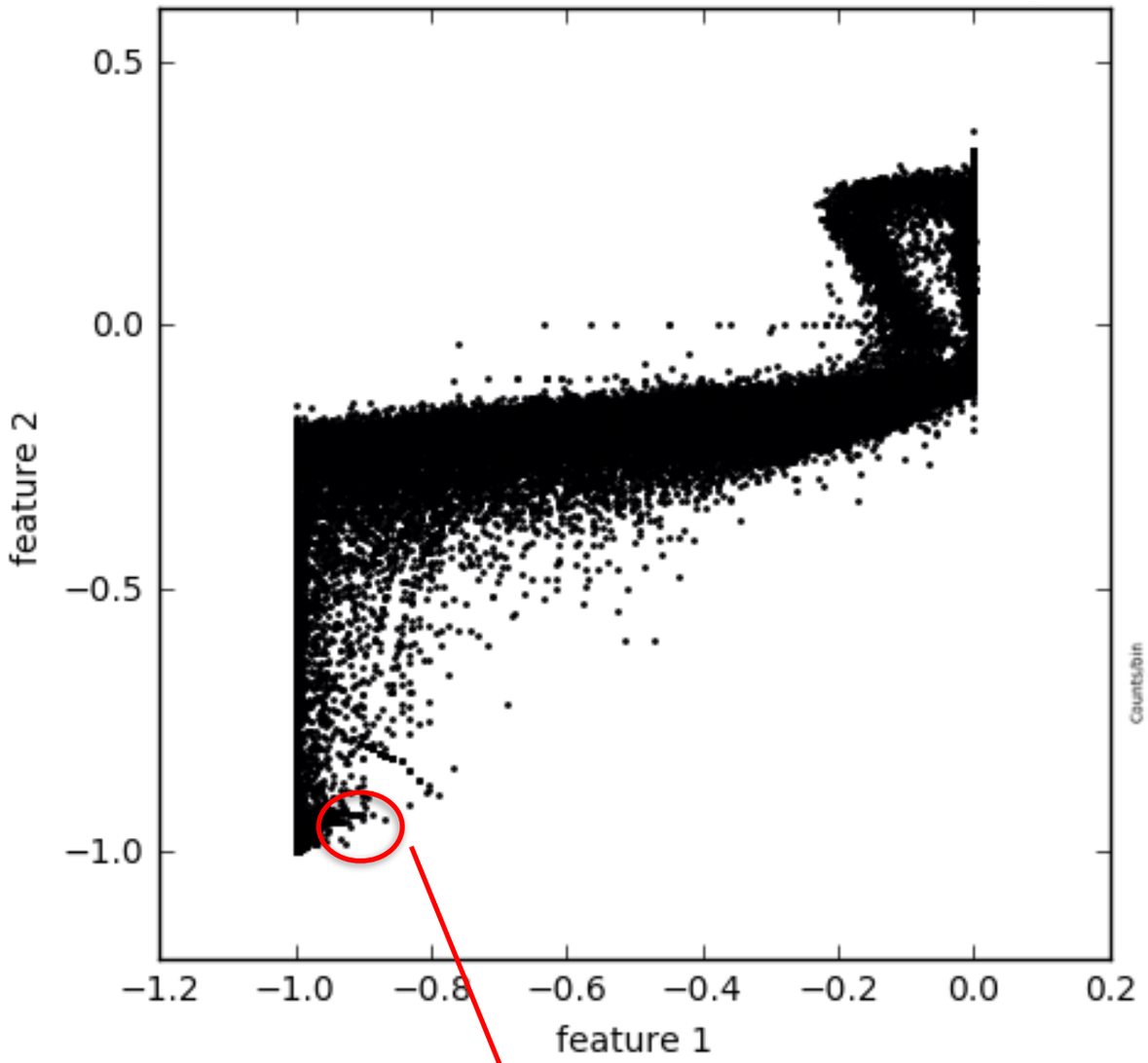


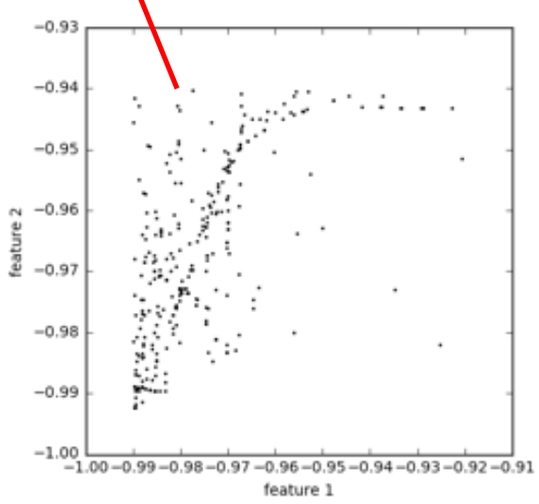
“Deep Analysis”

- Is this even useful?
- Note that if we have a machine that can do this, it must have learned something about the system. How can we extract this information?
  - Try to extract it from the outputs of specific neurons in the network

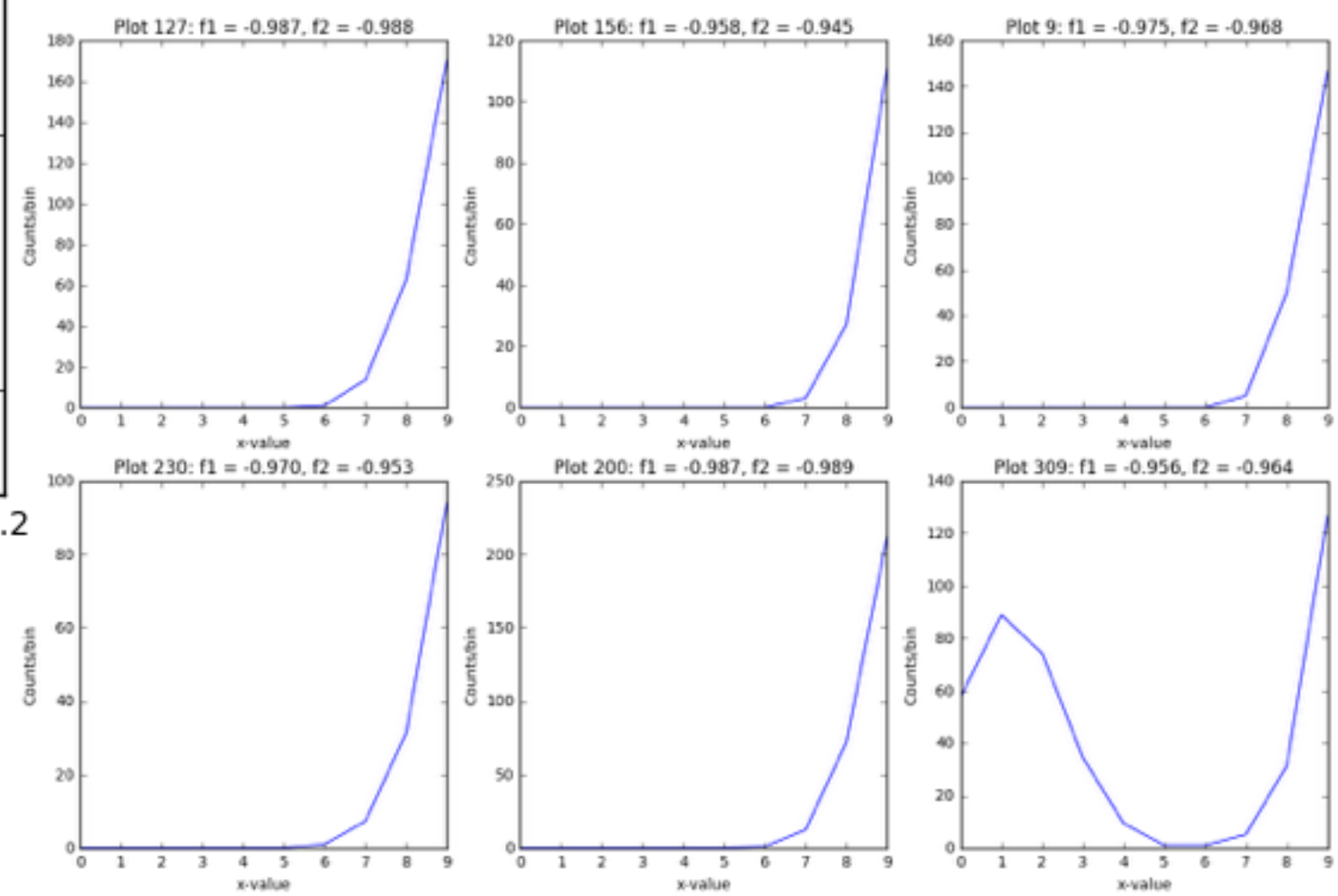


“Deep Analysis”

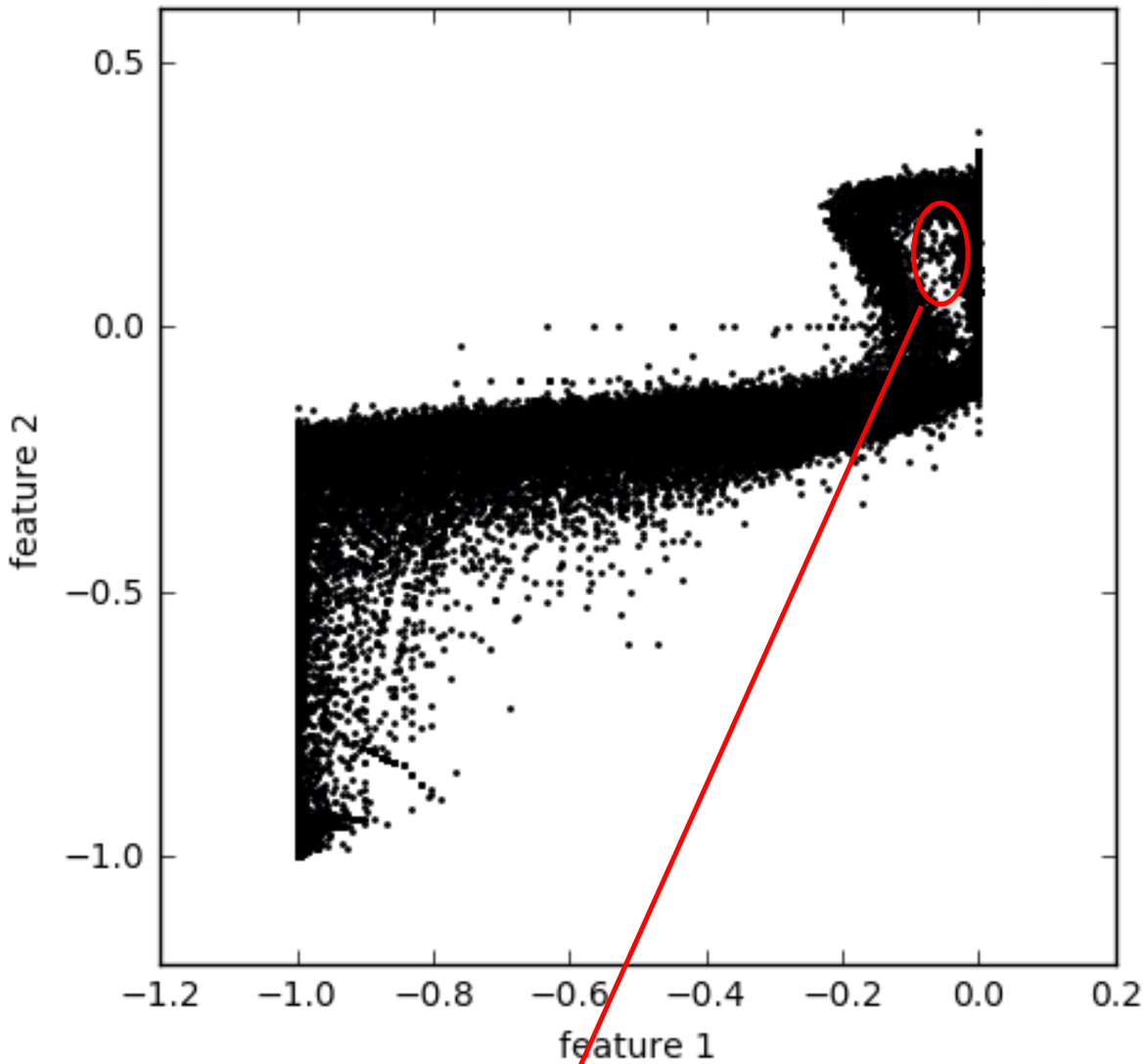


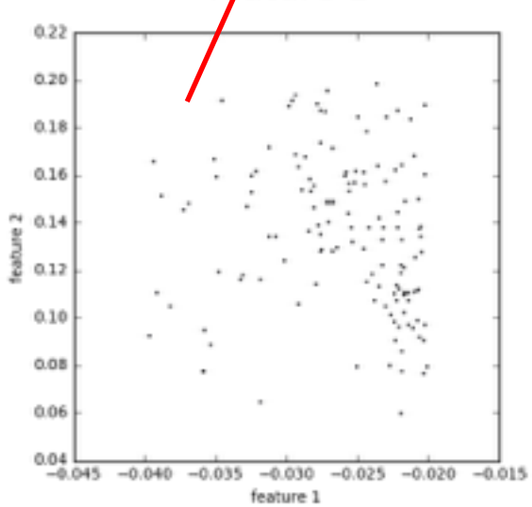


- Certain waveform segments fall in different locations in feature space



“Deep Analysis”





- Certain waveform segments fall in different locations in feature space