Dual-Readout Calorimetry Signal Analysis with Neural Networks

Murali Saravanan

Overview

- Dual-Readout Calorimetry Recap
- Current Results
  - Part I: Monte Carlo Simulations
  - Part II: Real Data
  - Part III: hls4ml
- Future Steps

The LHC and Calorimetry

- Calorimeters are responsible for energy measurement in ATLAS, CMS, and various particle detectors

- **Hadronic** and Electromagnetic Calorimetry

  protons, pions, and fragmenting quarks and gluons

Hadronic Calorimetry

- Measure em and non-em components
  - Huge fluctuations in em component (up to 40%) on event-by-event basis

- Different calorimeter materials respond to em and non-em differently. Use response curves but this works on average

Dual Readout Calorimetry Recap

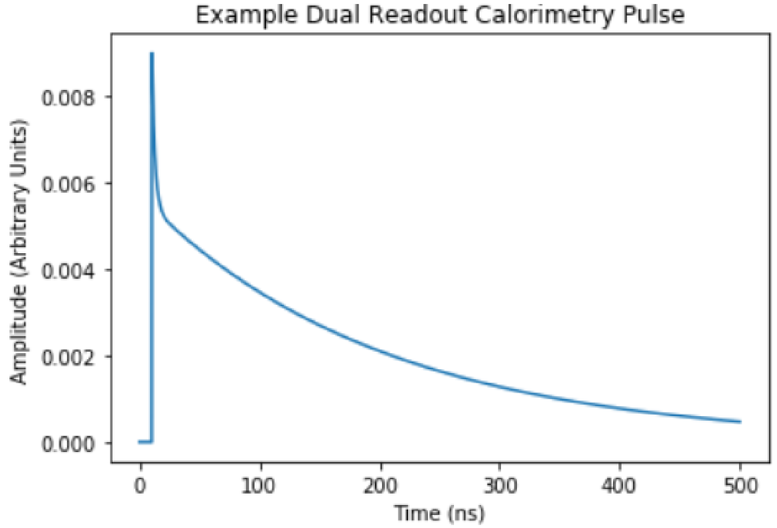
- Measure both Cerenkov and Scintillation radiation to achieve em energy resolution on event-by-event basis
  - Amount of Cerenkov/Scint light can be tweaked by materials and geometry of physical calorimeter

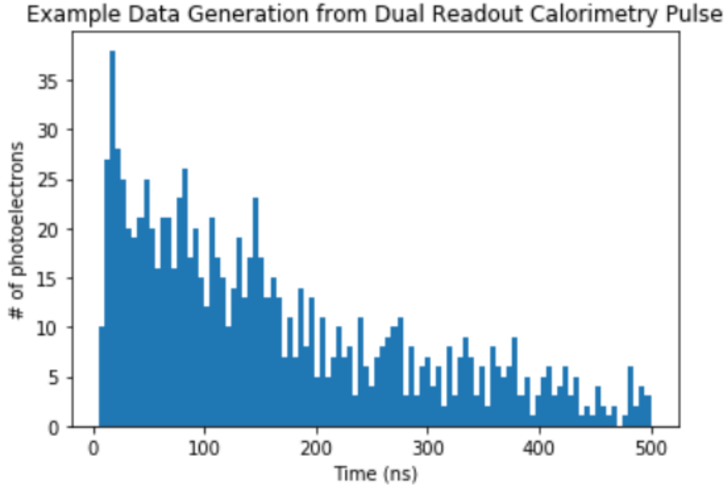
- Huge improvement for hadronic energy resolution, would be useful in future colliders (FCC, ILC, CLIC)

Single Channel Readout?

- Single Channel saves overhead
- Four parameters vary
  1. Ratio of Cerenkov/Scintillation Radiation (Area)
  2. Scintillation Decay Rate
  3. Photoelectron Count
  4. Digitizer Freq/Bin Number

Material Properties





Using NN

- NN is constant time
- Predict two parameters
  1. Cerenkov pulse area
  2. Scintillation pulse area

```
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
dense_5 (Dense)              (None, 10)                1010
_________________________________________________________________
dense_6 (Dense)              (None, 2)                 22
=================================================================
Total params: 1,032
Trainable params: 1,032
Non-trainable params: 0
_________________________________________________________________
None
```

Part 1: Monte Carlo Simulated Data

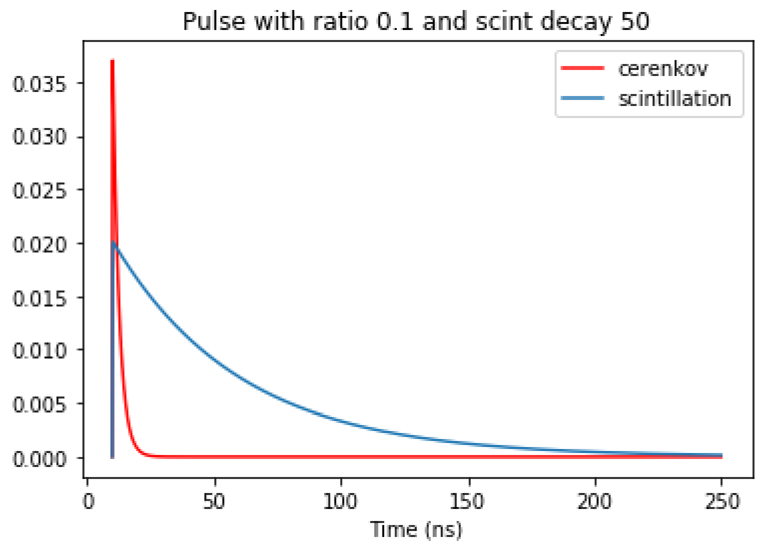
Monte-Carlo Generation of Data

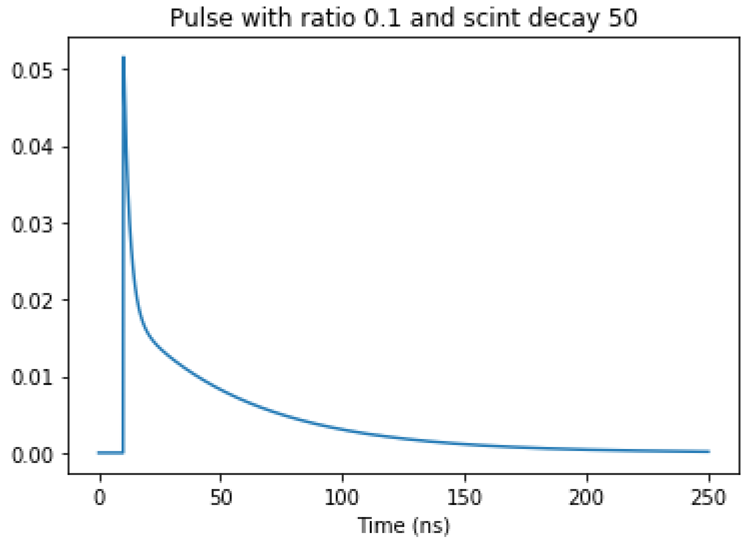
- Modify:
  - Ratio
    - 10-> 10x more area of Cerenkov than Scintillation
    - -10-> 10x more area of Scintillation than Cerenkov
  - Scintillation Decay
  - Photoelectron Count
  - Binning

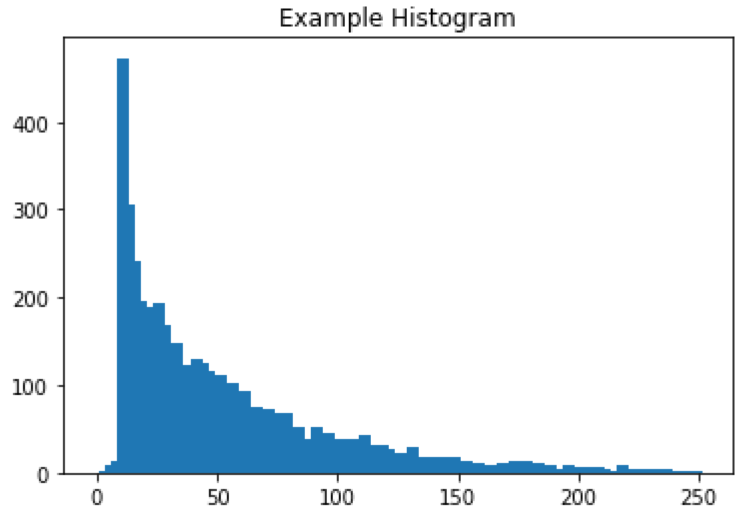
- Generate a library of histograms, varying the above parameters to create different experiments
- Create many events per experiment to create needed random fluctuations

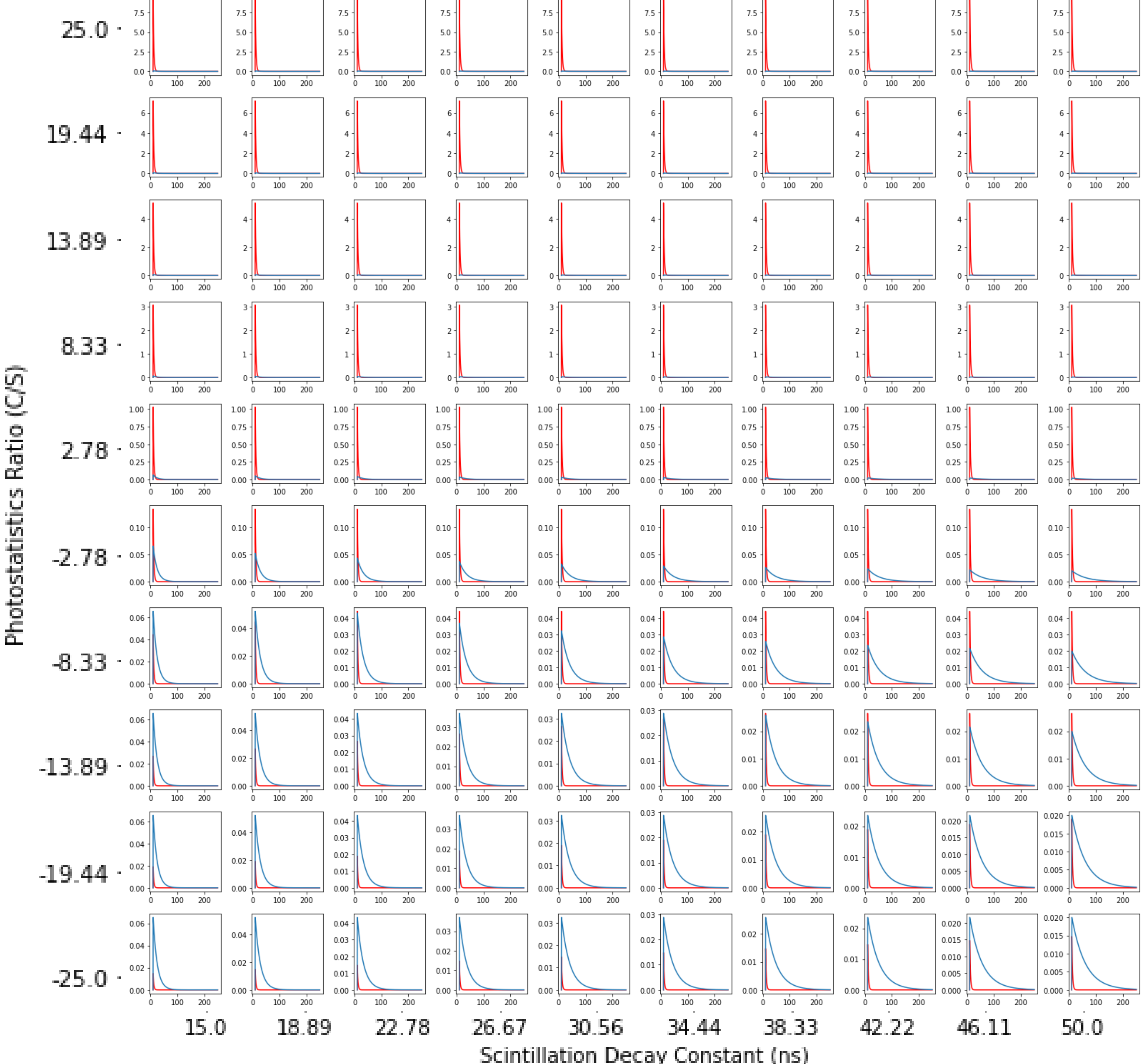
Example Pulse

- Vary Ratio between -25 and 25
- Vary Scint Decay from 15ns to 50ns
- 1k-5k photoelectrons
- 30, 100, 300 bins



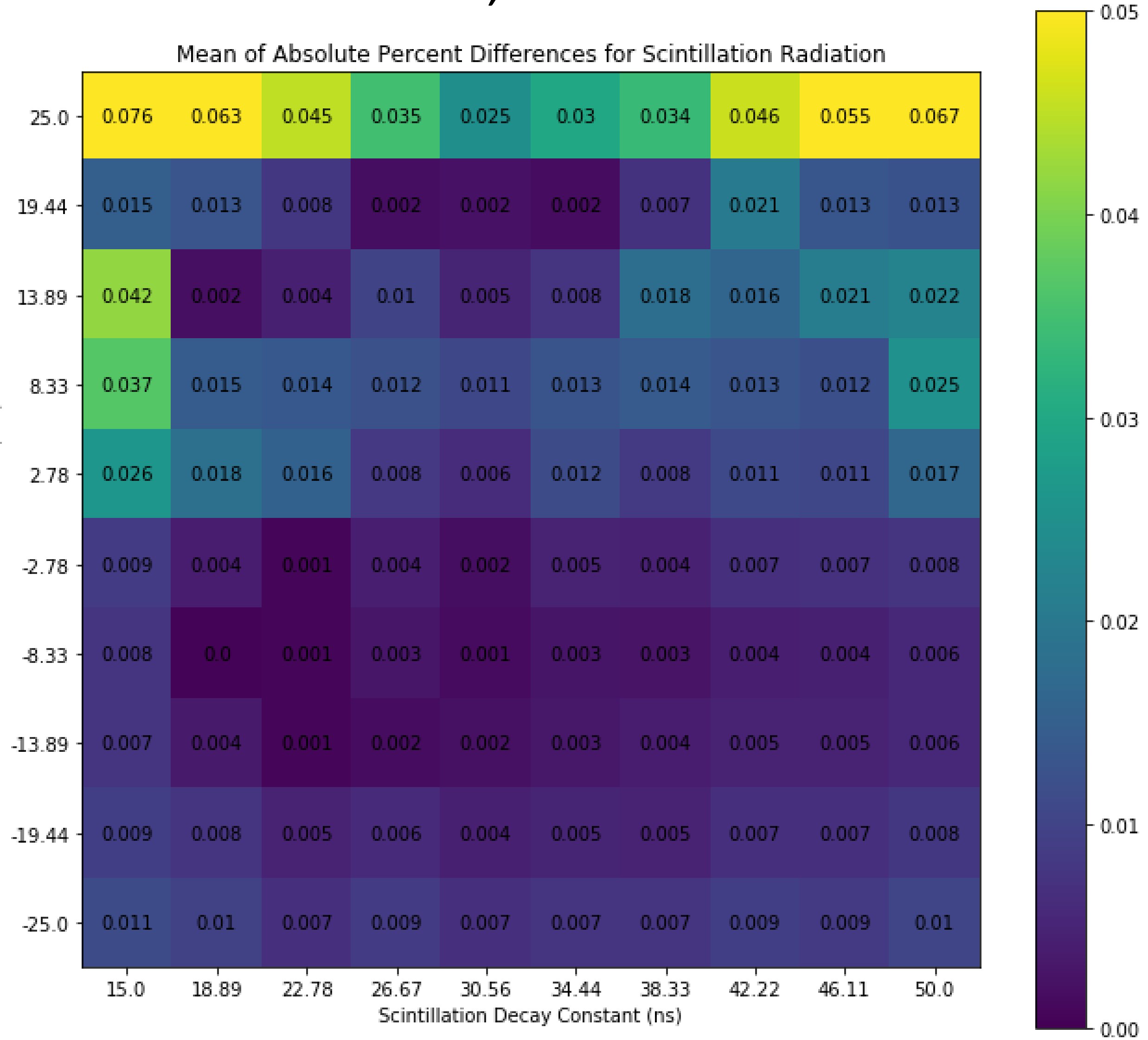


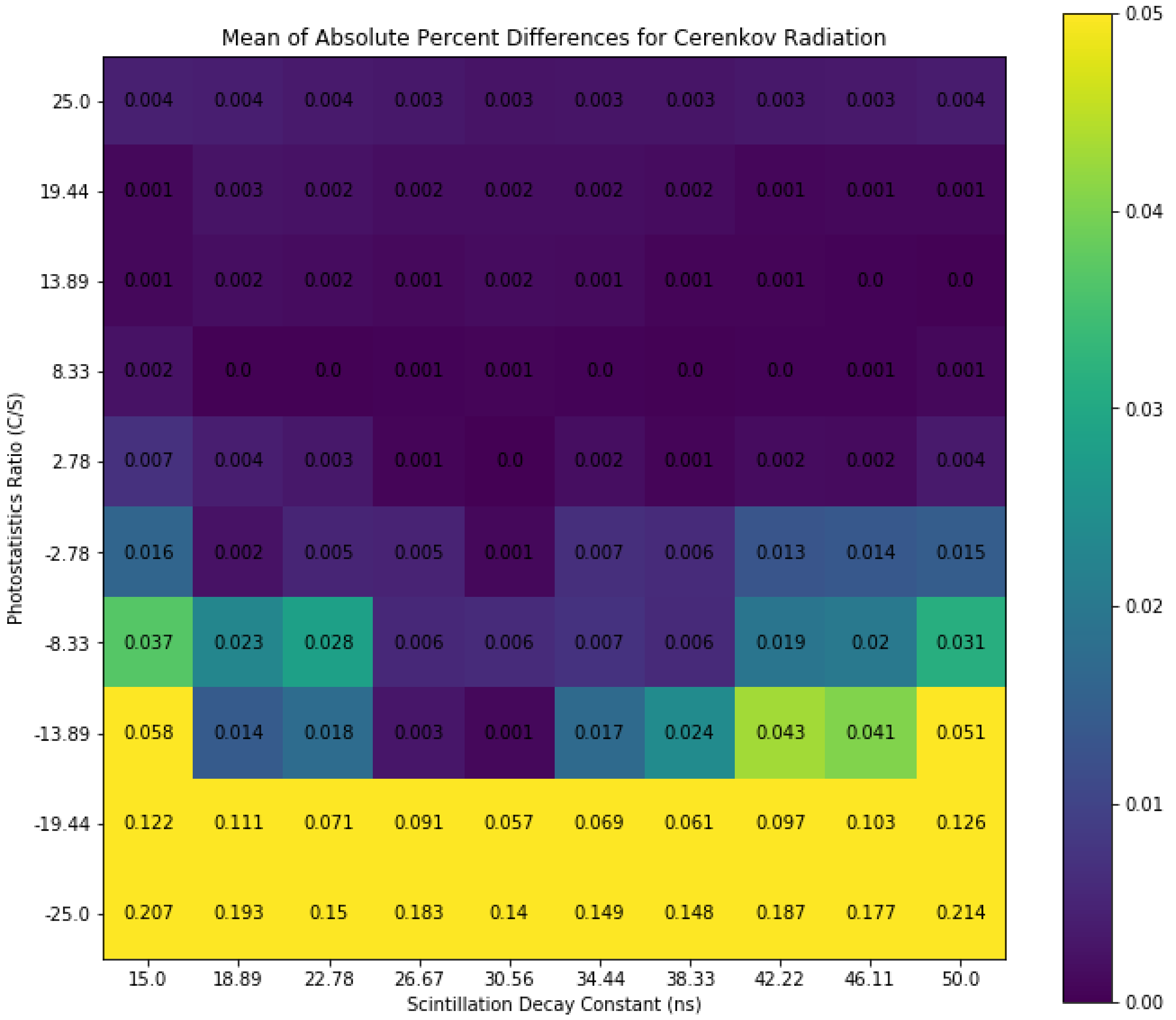




Photostatistics Ratio (C/S)
25.0
19.44
13.89
8.33
2.78
-2.78
-8.33
-13.89
-19.44
-25.0
15.0
18.89
22.78
26.67
30.56
34.44
38.33
42.22
46.11
50.0
Scintillation Decay Constant (ns)

- Vary digitizer freq and photostatistics
    - 30 bins, 5k





Part I Conclusions

- Major dependence is on ratio and not scintillation decay time
- Higher photoelectron count gives better results
  - Gradual change over 1k-5k photoelectrons range (no tipping point)
- Higher digitizer freq ≠ better prediction
  - Low freq hides the effects of fluctuations
- Scintillation prediction is much more stable than Cerenkov prediction

Part 2: Real Data

Dual Cherenkov and Scintillation Response to High-Energy Electrons of Rare-Earth-Doped Silica Fibers

Francesca Cova,[1] Marco T. Lucchini,[2,3,*] Kristof Pauwels,[4,3,†] Etiennette Auffray,[3] Norberto Chiodini,[1] Mauro Fasoli,[1] and Anna Vedda[1]

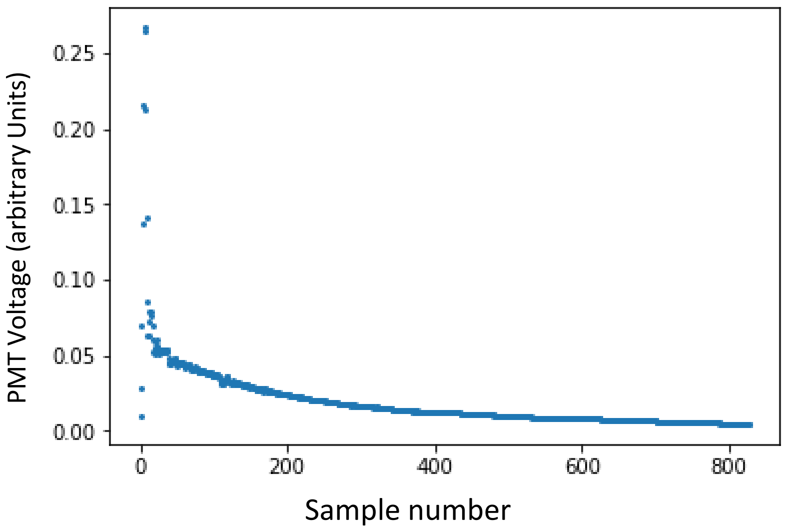
[1] *Department of Materials Science, University of Milano-Bicocca, Via Cozzi 55, 20125 Milan, Italy*
[2] *Princeton University, Princeton, New Jersey 08544, USA*
[3] *EP-CMX Department, CERN, 1211 Geneva 23, Switzerland*
[4] *Department of Physics, University of Milano-Bicocca, Piazza della Scienza 3, 20125 Milan, Italy*

Real Data

- response of Ce-doped silica fibers exposed to electrons in the 20–200-GeV
- What does a realistic pulse look like? What is the actual ratio of C/S? Expected photoelectron count?

Average Pulse

200 GeV Pulse



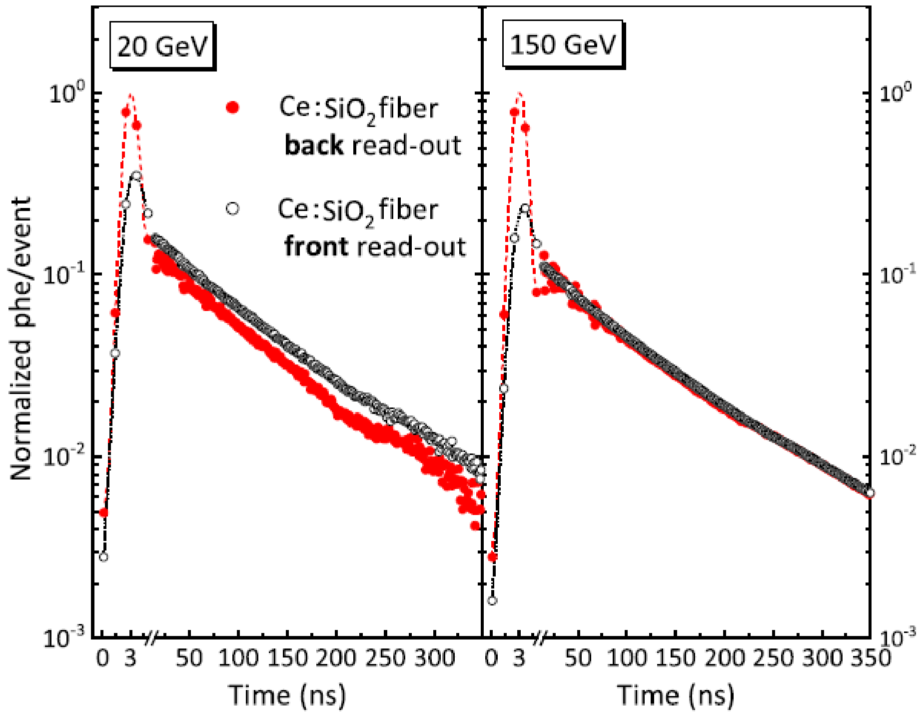
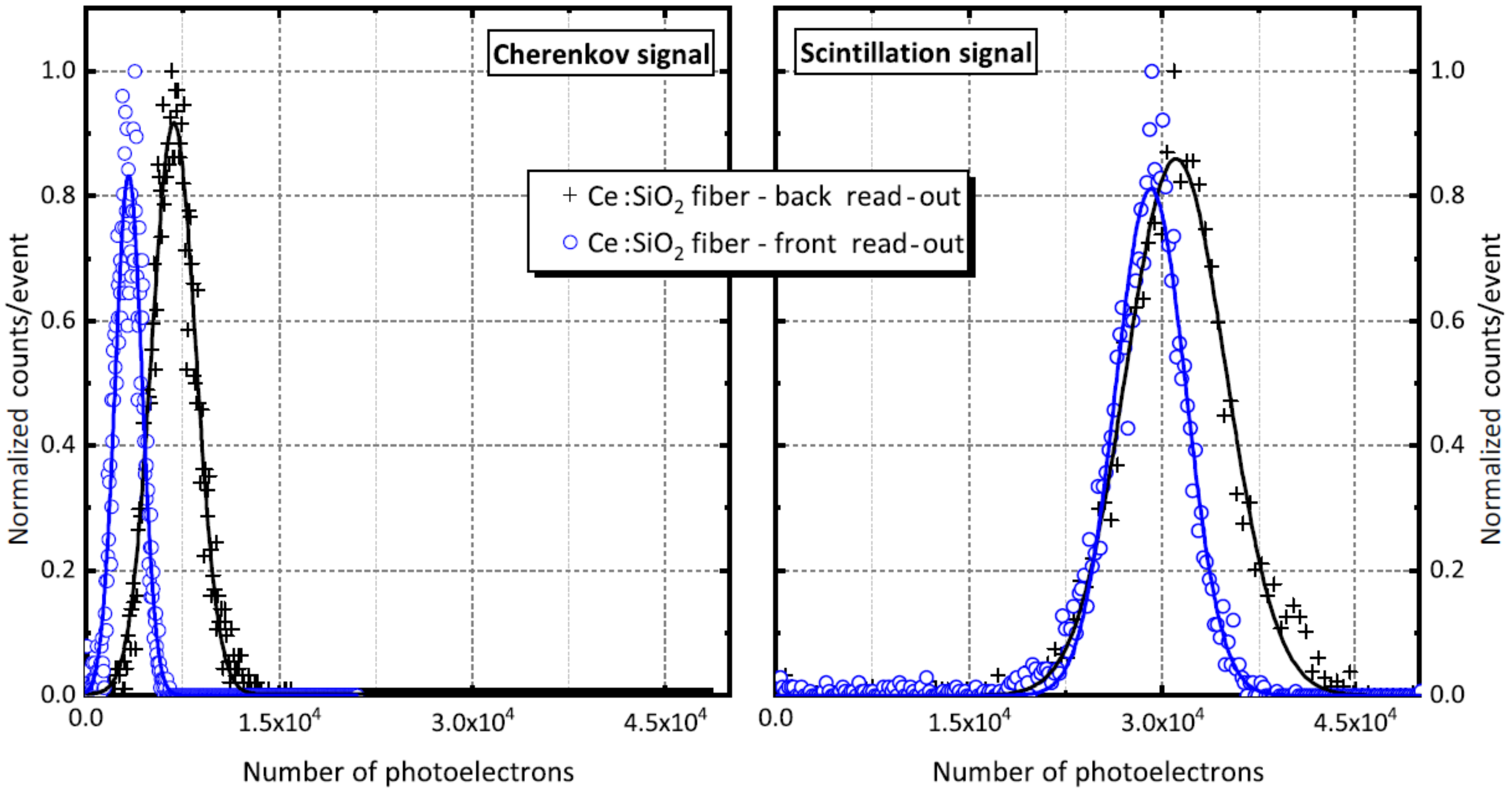


FIG. 6. Average pulse shape normalized to the number of events for Ce:SiO_2 fibers for a 20-GeV electron beam (left panel) and a 150-GeV electron beam (right panel). Back and front read-out are compared. The dashed lines are guides for the eye. phe, photoelectrons.

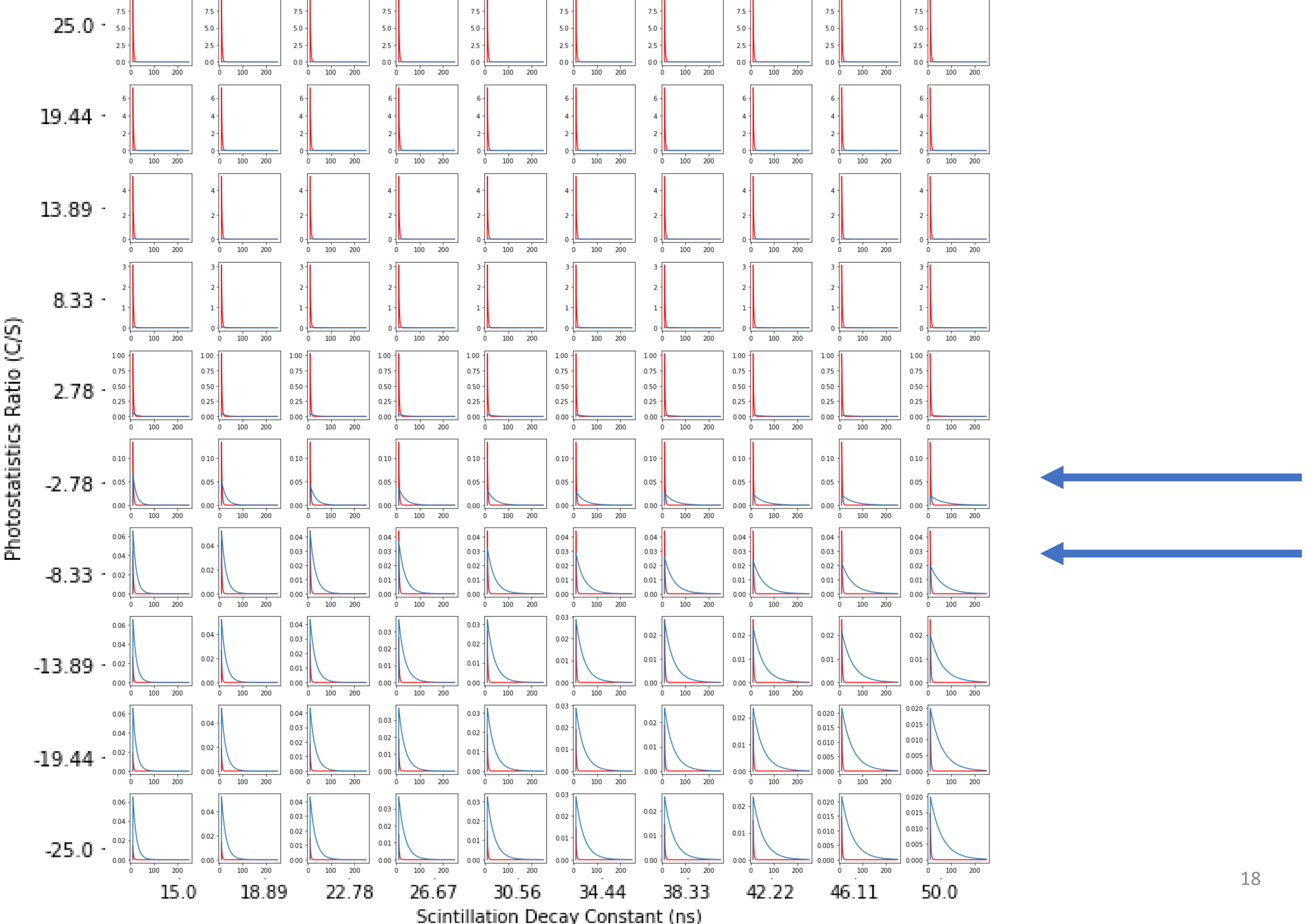
[1] Cova, Francesca, et al. *Physical Review Applied* 11.2 (2019): 024036.

Photoelectron Count

- 7.5k for Cerenkov, 30k for Scintillation (ratio=-4)



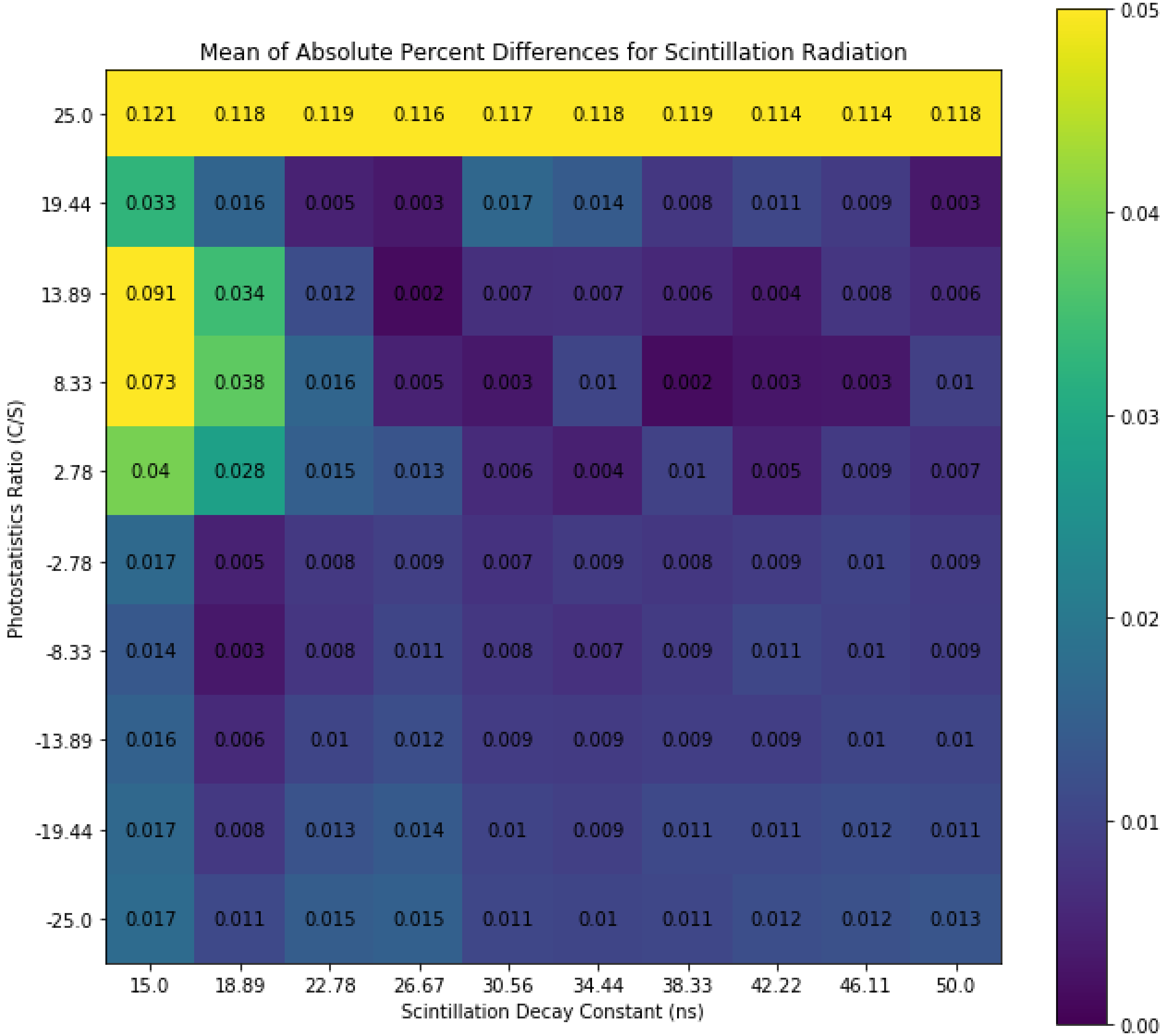
[1] Cova, Francesca, et al. *Physical Review Applied* 11.2 (2019): 024036.



Photostatistics Ratio (C/S)
25.0
19.44
13.89
8.33
2.78
-2.78
-8.33
-13.89
-19.44
-25.0
15.0
18.89
22.78
26.67
30.56
34.44
38.33
42.22
46.11
50.0
Scintillation Decay Constant (ns)

Monte Carlo results in this area

- With 5,000 photoelectrons and 100 bins, possibility of less than 1% error



Real Data NN

- In progress!
- Need to consistently clean data

Part 3: Implementation on FPGAs

Intro to FPGAs

- Multiplier Units
  - Does arithmetic
- Flip Flops
  - Short term memory registers

- Wire algorithms into board
- High Level Synthesis (HLS) is the language that Xilinx FPGAs use

hls4ml

- Takes NN and creates HLS implantation that provides resource usage estimates

- Up to 6,000 parallel operations=# of multiplication units=max # of nodes and weights in NN if we aim for one clock cycle per analysis

- Can trade latency for lower resource usage

[2] J.M. Duarte *et al.*, arXiv:1804.06913, April 2018.

Preliminary Results

- 100 bin Model with 5ns goal

Synthesis Report for 'testModel'

General Information

Date:	Tue Apr 9 15:50:37 2019
Version:	2017.2 (Build 1909853 on Thu Jun 15 18:55:24 MDT 2017)
Project:	testModel_prj
Solution:	solution1
Product family:	kintexu
Target device:	xcku115-flvb2104-2-i

Performance Estimates

Timing (ns)

Summary

Clock	Target	Estimated	Uncertainty
ap_clk	5.00	4.37	0.63

Latency (clock cycles)

Summary

Latency		Interval		
min	max	min	max	Type
4	4	1	1	function

Detail

Instance

Loop

Utilization Estimates

Preliminary Results

- hls4ml in general provides an overestimate of resource usage

Utilization Estimates

Summary

Name	BRAM_18K	DSP48E	FF	LUT
DSP	-	-	-	-
Expression	-	-	0	6
FIFO	-	-	-	-
Instance	-	718	19902	16648
Memory	-	-	-	-
Multiplexer	-	-	-	36
Register	-	-	1766	-
Total	0	718	21668	16690
Available	4320	5520	1326720	663360
Available SLR	2160	2760	663360	331680
Utilization (%)	0	13	1	2
Utilization SLR (%)	0	26	3	5

- Can prune unnecessary nodes from intermediate layer to decrease resource usage

Project Conclusions

- Dual Readout Calorimetry is promising for improving resolution of em-portion of hadronic energy on an event-by-event basis
- Neural Networks can be implemented into FPGAs for low latency inference in **single channel** dual readout calo
  - Small networks are adequate
- Simulated performance of NNs can be used to inform the geometry/construction of the calorimeters and the associated hardware

Project Conclusions

Single Channel Dual Readout Calorimetry can work thanks to Neural Networks

Future Steps

- Produce NN results on real data
- Look at implementation on real FPGA

Questions?