Electron Identification with Deep Neural Networks

Dominique Godin
Kazuya Mochizuki
Emre Onur Kahya
Arthur Boisvert

Jean-François Arguin
Alain Tapp
Otilia Ducu
Julien Maurer
Ioan-Mihail Dinu

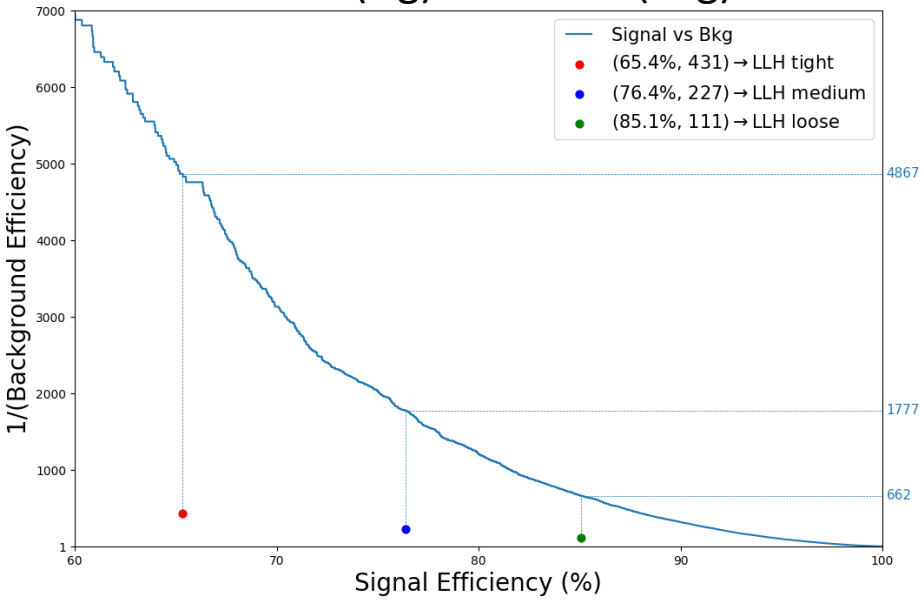
ATLAS EXPERIMENT

Université de Montréal

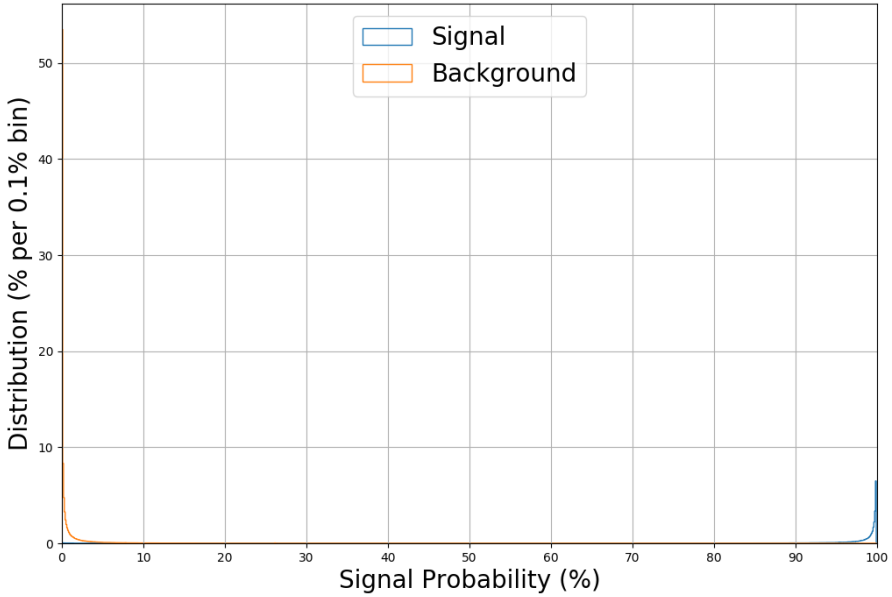
2-Class Preliminary Results

- MC sample: 9 million training and 1 million testing electrons
- Neural network: ≈ 400 000 training parameters

Test sample ROC curve:
class 0 (sig) vs class 1 (bkg)



Separation of test sample
distributions after training



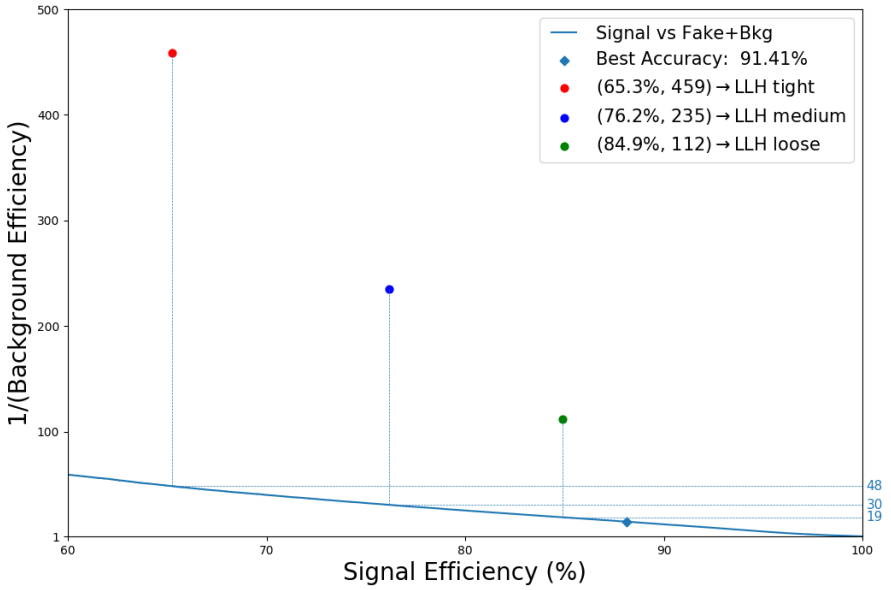
CLASS DISTRIBUTION AND ACCURACY:

CLASS	TRAIN(%)	TEST(%)	ACC.(%)
class 0	36.73	36.72	96.34
class 1	63.27	63.28	98.39

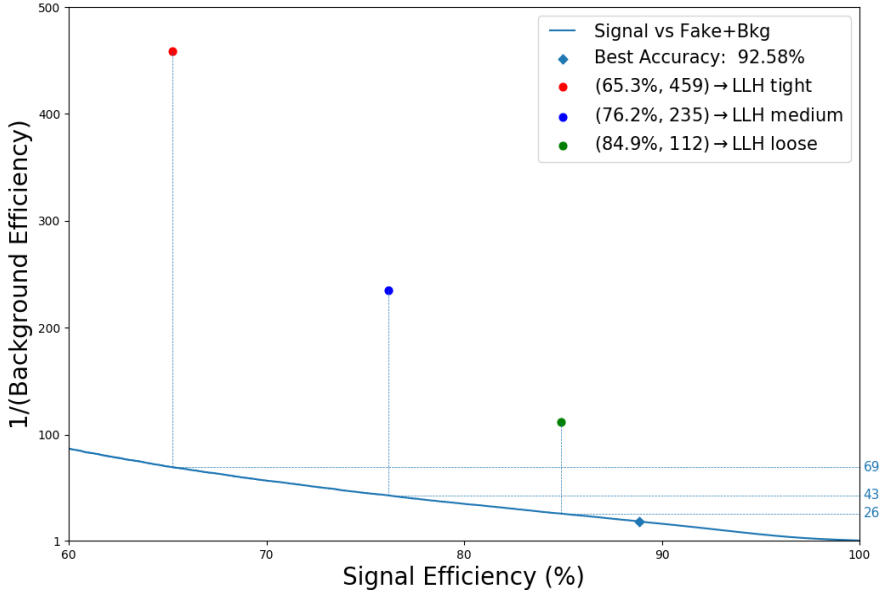
2-Class Preliminary Results with 1st EM Layer

- Trainings with only 1st EM layer show that more information can be learned from NN with a finer resolution (i.e. 56x11)

Test sample ROC curve
(EM 1st layer 7x11)



Test sample ROC curve
(EM 1st layer 56x11)



- MC sample: 5m electrons (4.5m for training, 0.5m for testing)
- Neural network: ≈ 400 000 training parameters

5-Class Preliminary Results

- Classes definitions:

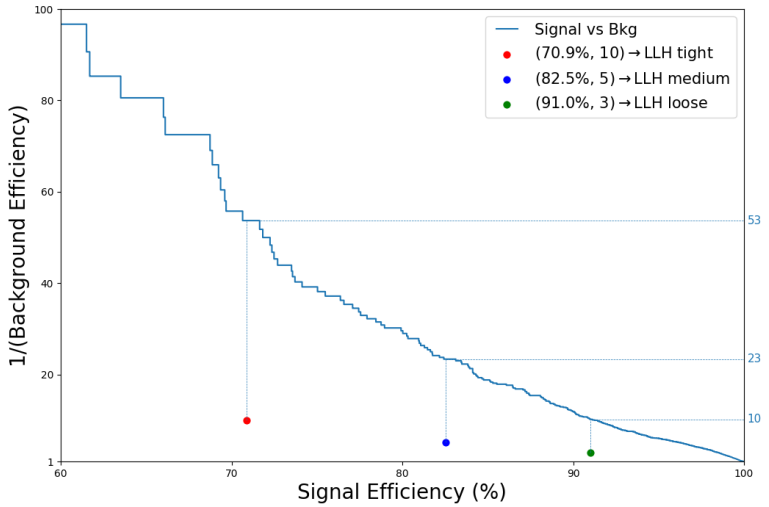
```
truthIFF== 0: truthTypeMod=4 # Unknown
truthIFF== 1: truthTypeMod=2 # KnownUnknown
truthIFF== 2: truthTypeMod=0 # IsoElectron
truthIFF== 3: truthTypeMod=1 # ChargeFlipIsoElectron
truthIFF== 4: truthTypeMod=4 # PromptMuon
truthIFF== 5: truthTypeMod=4 # PromptPhotonConversior
truthIFF== 6: truthTypeMod=4 # ElectronFromMuon
truthIFF== 7: truthTypeMod=3 # TauDecay
truthIFF== 8: truthTypeMod=3 # BHadronDecay
truthIFF== 9: truthTypeMod=3 # CHadronDecay
truthIFF==10: truthTypeMod=2 # LightFlavorDecay
```

- MC sample: 10m electrons (9m for training, 1m for testing)
- Neural network: ≈ 400 000 training parameters
- Confusion matrix after training:

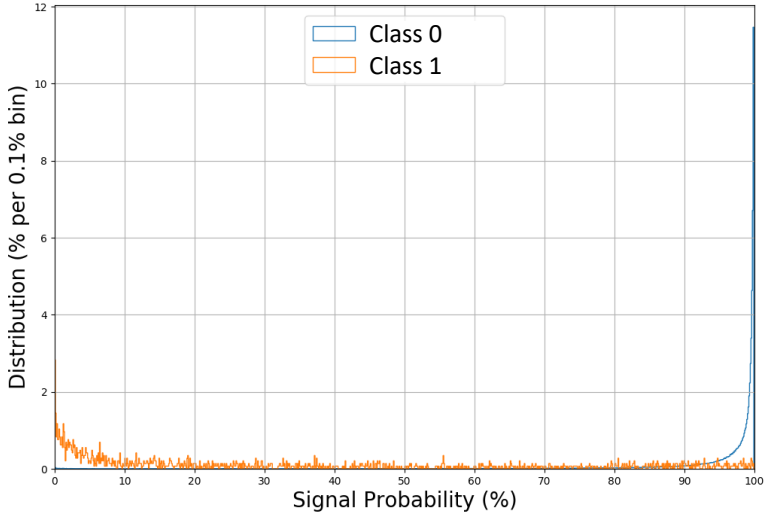
CLASS DISTRIBUTIONS (%)			TEST PREDICTIONS (%)				
CLASS #	TRAIN	TEST	CLASS 0	CLASS 1	CLASS 2	CLASS 3	CLASS 4
CLASS 0	33.27	33.28	98.02	31.72	0.43	44.16	15.61
CLASS 1	0.15	0.14	0.05	31.10	0.00	0.07	0.32
CLASS 2	62.72	62.73	1.02	13.59	98.93	23.06	16.19
CLASS 3	1.36	1.35	0.46	0.97	0.18	31.95	0.42
CLASS 4	2.50	2.50	0.45	22.62	0.46	0.75	67.46

5-Class Preliminary Results

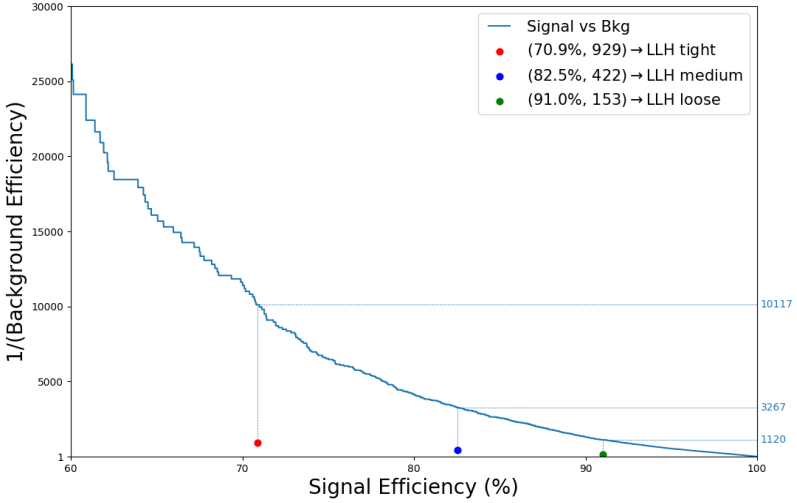
Test sample ROC curve (class 0 vs class 1)



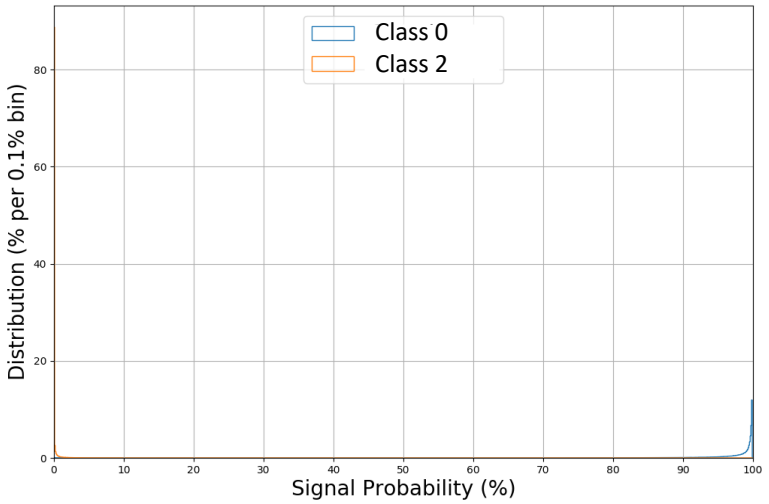
Separation of test sample distributions (class 0 vs class1)



Test sample ROC curve (class 0 vs class 2)

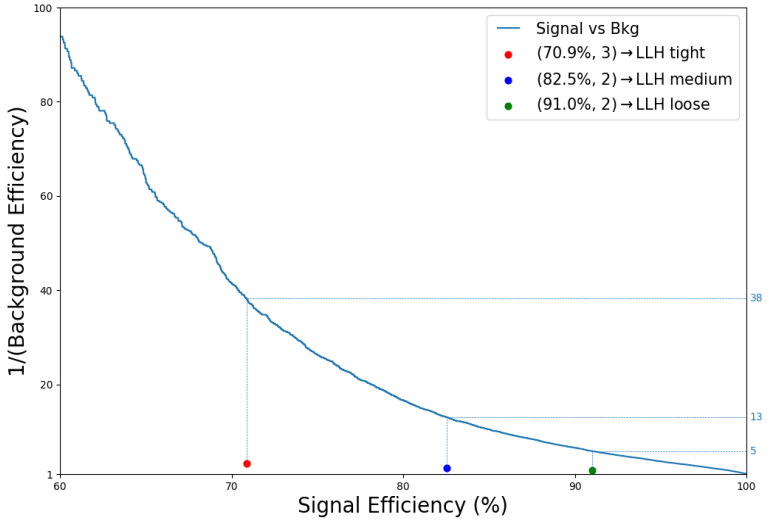


Separation of test sample distributions (class 0 vs class 2)

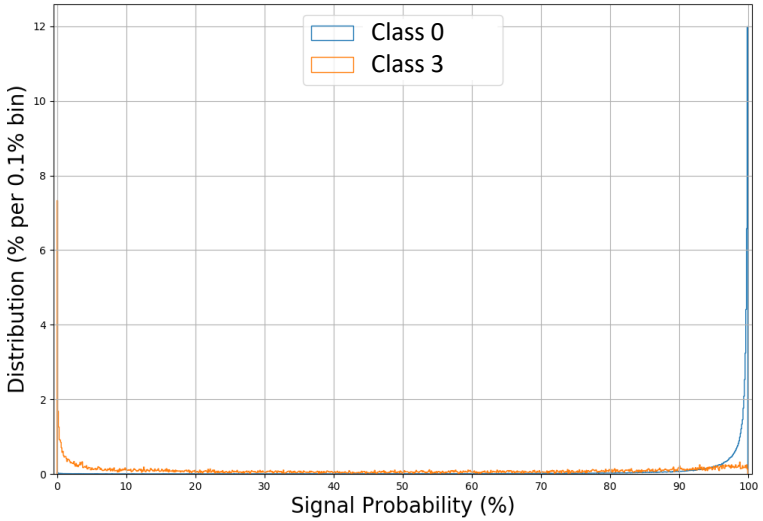


5-Class Preliminary Results

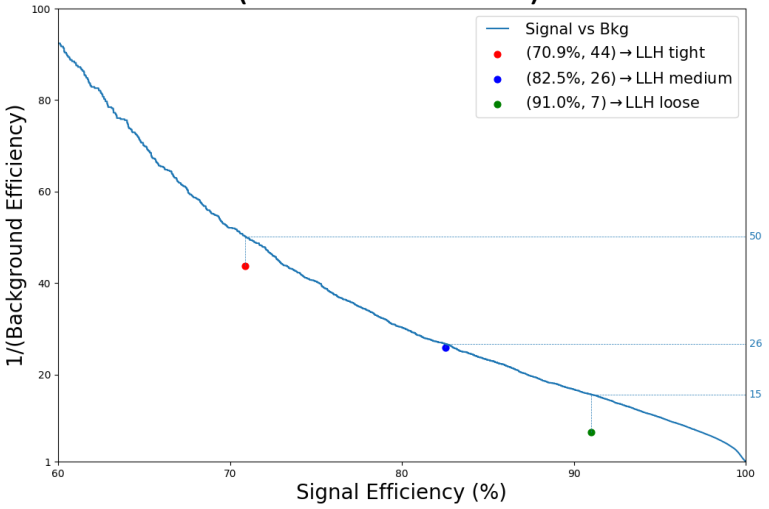
Test sample ROC curve
(class 0 vs class 3)



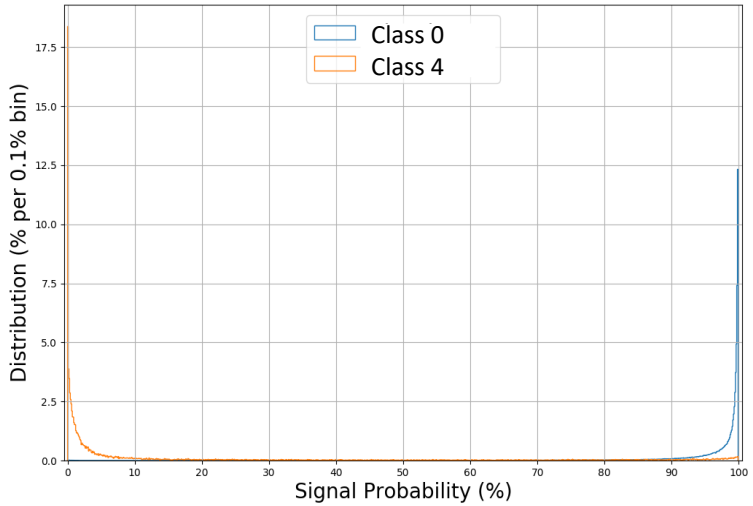
Separation of test sample
distributions (class 0 vs class 3)



Test sample ROC curve
(class 0 vs class 4)



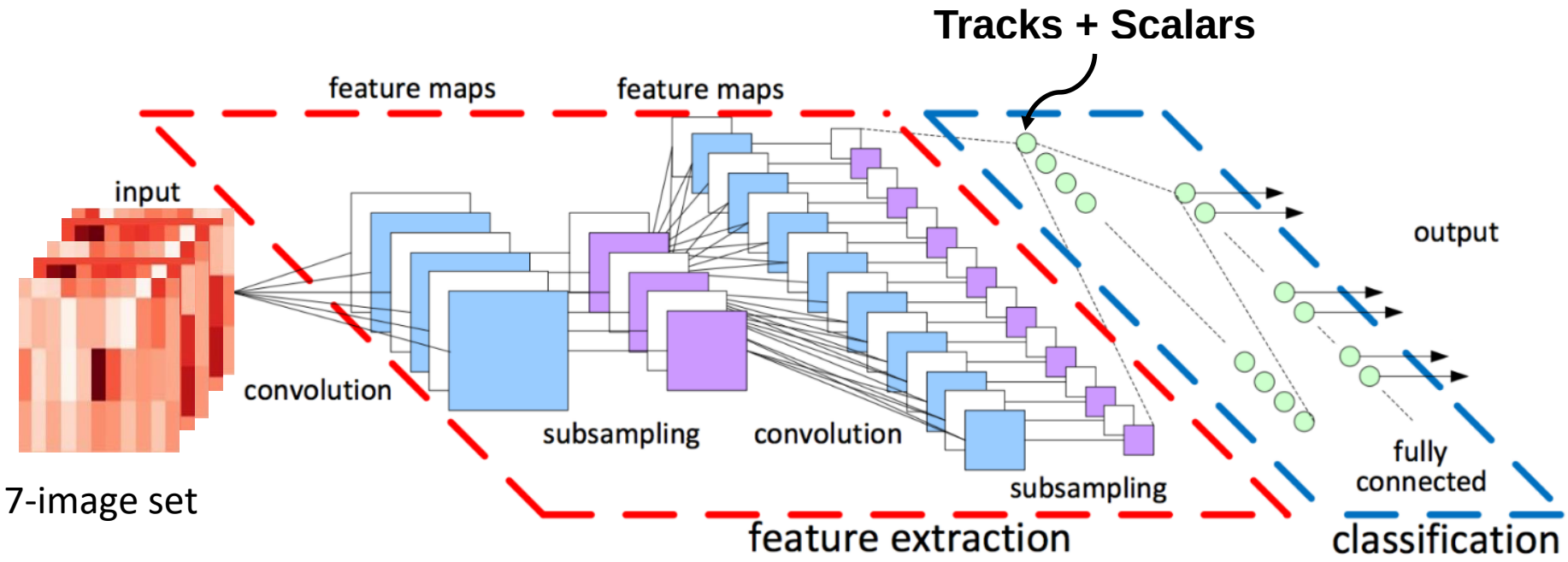
Separation of test sample
distributions (class 0 vs class 4)



Backup Material

Preliminary NN Architecture

- Multichannel CNN with calorimeter images + tracks and scalars information
- The 7-image sets are concatenated in volume images (7-channel images)
- Tracks and scalars concatenated in the FC layer

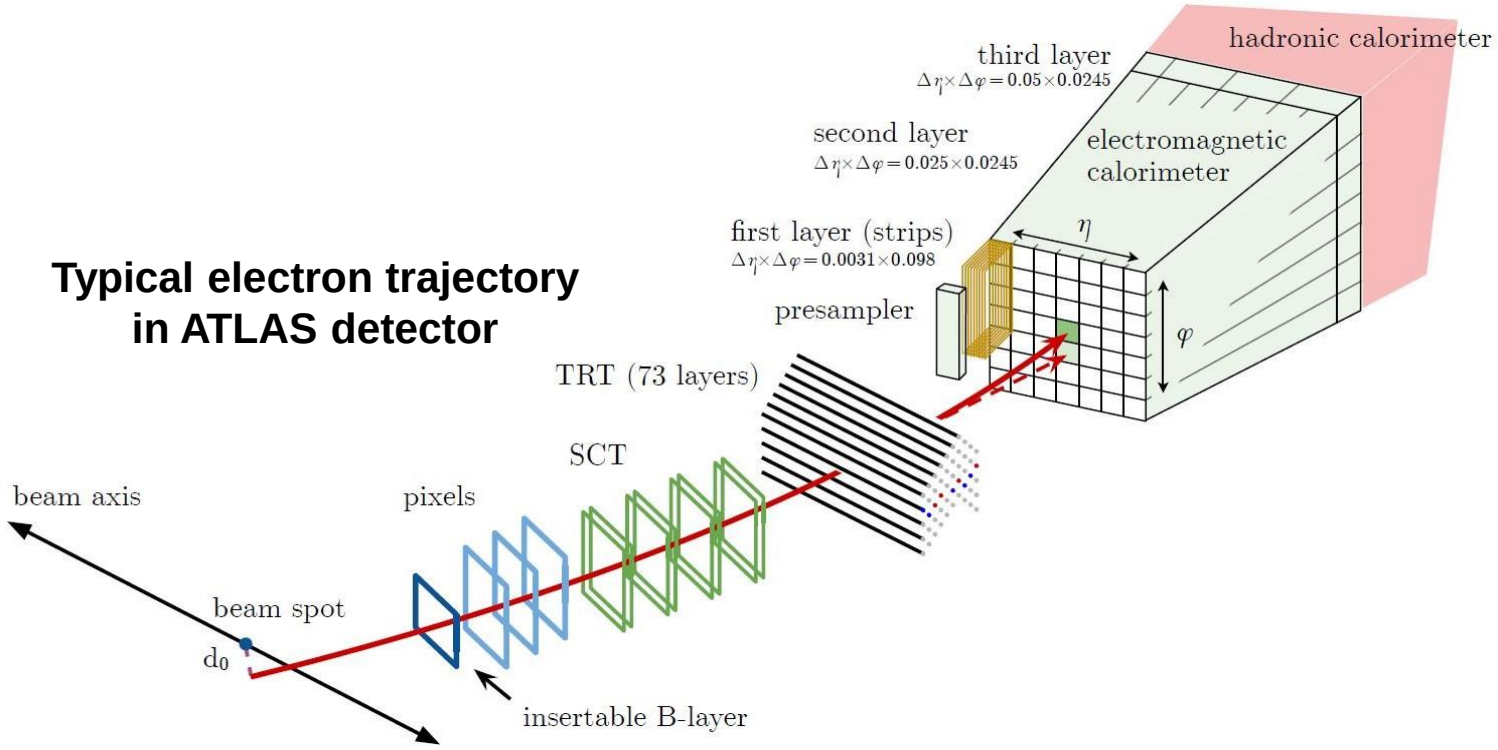


Electron Data from ATLAS Detector

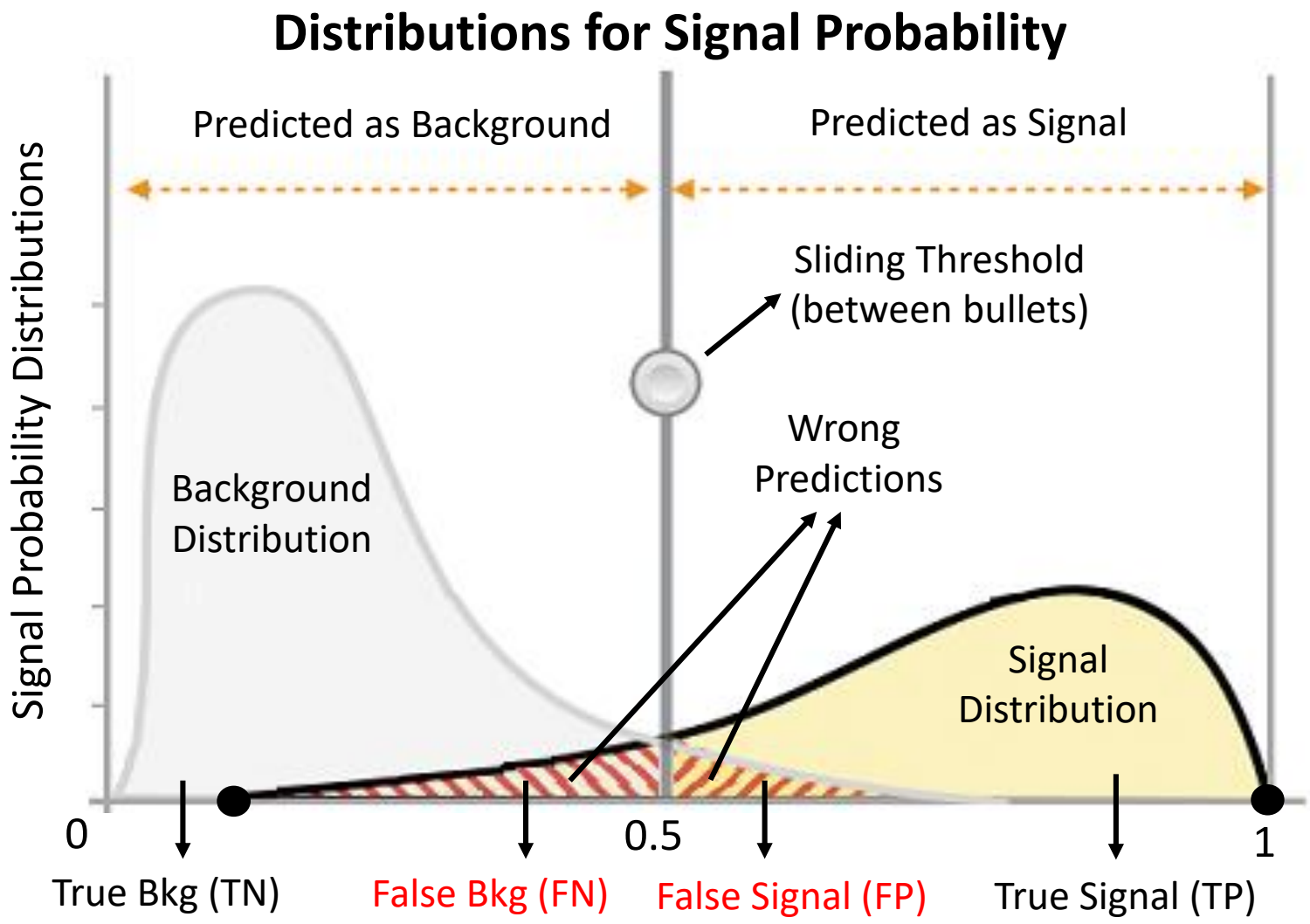
Available MC data for each electron

- 4 images from pre-sampler and EM calorimeter (56 x 11 in η x φ space)
- 3 images from hadronic calorimeter (7 x 11 in η x φ space)
- Up to 15 track candidates with (e_frac, d_eta, d_phi, d_0) information
- Scalars information such as P_t, φ, d_0, shower variables, LLH inputs, etc.

Typical electron trajectory in ATLAS detector



Probability Distributions / ROC Curves



$$\text{Background Rejection} = \frac{TN}{TN+FP} \quad \text{vs} \quad \text{Signal Efficiency} = \frac{TP}{TP+FN}$$

Project Group Members

U. Montreal
- Jean-François Arguin
- Otilia Ducu
- Dominique Godin
- Kazuya Mochizuki

Université de Montréal

Mila

Montreal Institute for Learning Algorithms
- Emre Kahya
- Alain Tapp: Prof.

IFIN-HH

IFIN-HH Romania
- Ioan Dinu
- Julien Maurer

IVADO

- This project is largely supported by IVADO
- Thank you to the Copenhagen group for providing us with training samples