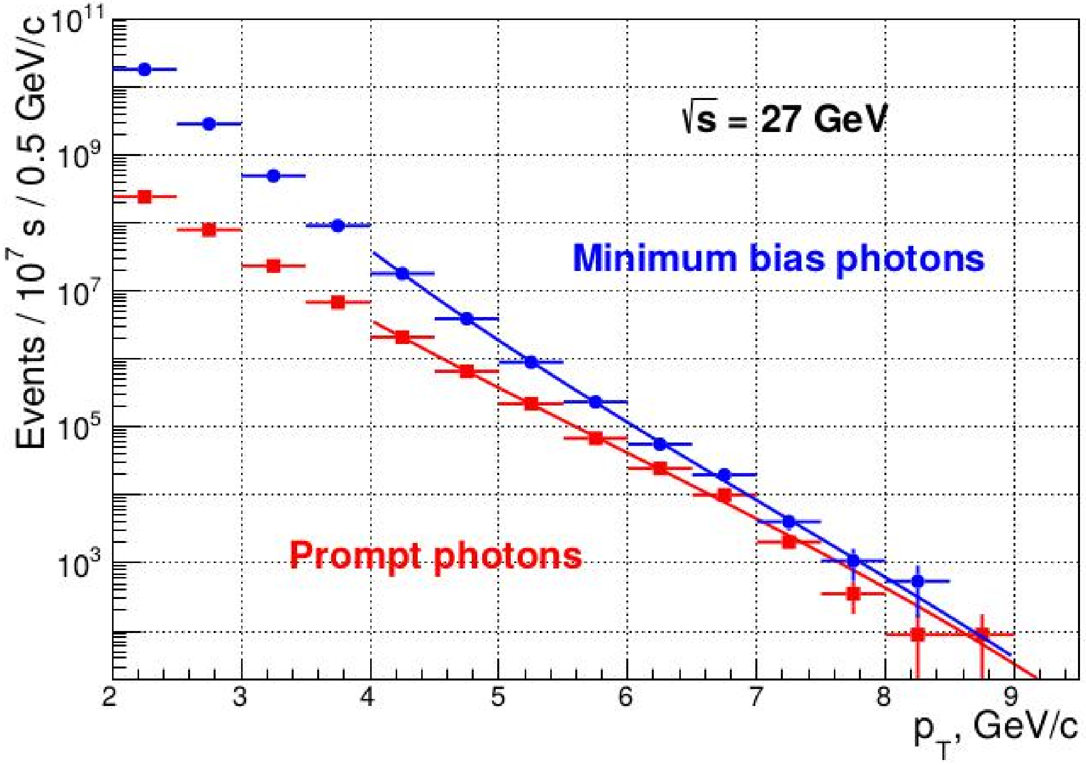
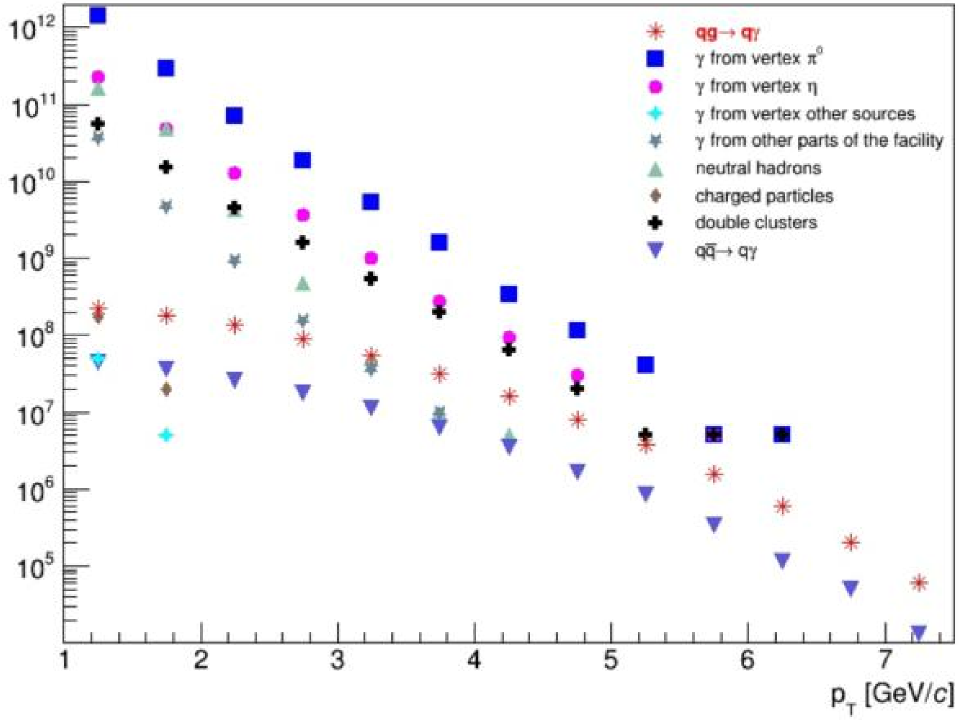
Performance of Machine Learning Techniques for $\pi0/\gamma$ Separation in SPD ECAL under Conditions Approximating Real Data

Andrei Maltsev
JINR, Dubna

SPD Physics & MC meeting
08.09.2021

Measurements with prompt photons





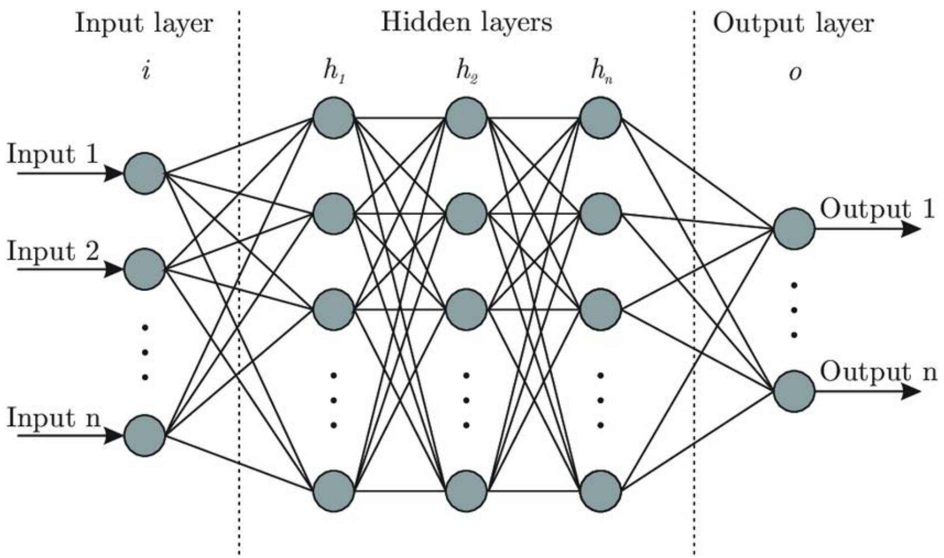
$$N_{prompt} = N_{\gamma} - k \times N_{\pi^0}$$

k: ratio of undetected π^0/η/... decays (from MC) giving “fake” propmpt photons

Beam polarization via π^0 production assymetry: better statistics

Attempt at using a more complex NN

Inspired by the work of Dimitrije Maletic (thanks!) and https://cds.cern.ch/record/2042173



$O_i = f(W_{i0} + \sum_{j=1}^{N} W_{ij} O_j)$ → weighted sum + bias for each node

- f: **ReLU** $f(x) = \begin{cases} 0 & \text{for } x < 0 \\ x & \text{for } x \geq 0 \end{cases}$ sigmoid for output: $f(x) = \frac{1}{1+e^{-x}}$

- **Dropout** (p=0.1),
- **batchnorm** for each layer (before activation)

- Binary cross entropy loss **(BCE)**:

$$H_p(q) = -\frac{1}{N}\sum_{i=1}^{N} y_i \cdot log(p(y_i)) + (1 - y_i) \cdot log(1 - p(y_i))$$

- Optimizer: **Adam**

(stochastic gradient descent +
adaptive moment estimation)
(lr = 0.001, β_1 = 0.9, β_2 = 0.999, ε=1e-8

2 hidden layers, 64 neurons each

Inputs

θ/φ moments:

$$|x_{cog}|_{25} = \left| \frac{\sum_{i=1}^{25} E_i X_i^{rel}}{S_{25}} \right|$$

$$|y_{cog}|_{25} = \left| \frac{\sum_{i=1}^{25} E_i Y_i^{rel}}{S_{25}} \right|$$

Correlation:

$$r2 =< r^2 >= S_{XX} + S_{YY} = \frac{\sum_{i=1}^{N} e_i((x_i - x_c)^2 + (y_i - y_c)^2)}{\sum_{i=1}^{N} e_i}$$

$$S_{XX} = \frac{\sum_{i=1}^{N} e_i(x_i - x_c)^2}{\sum_{i=1}^{N} e_i}, \quad S_{YY} = \frac{\sum_{i=1}^{N} e_i(y_i - y_c)^2}{\sum_{i=1}^{N} e_i},$$

$$S_{XY} = S_{YX} = \frac{\sum_{i=1}^{N} e_i(x_i - x_c)(y_i - y_c)}{\sum_{i=1}^{N} e_i},$$

Importance of tails:

$$r2r4 = 1 - \frac{< r^2 >^2}{< r^4 >}$$

Shape variable:

$$\kappa = \sqrt{1 - 4\frac{S_{XX}S_{YY} - S_{XY}^2}{(S_{XX} + S_{YY})^2}} = \sqrt{1 - 4\frac{\det S}{\mathrm{Tr}^2 S}}$$

Energy distribution

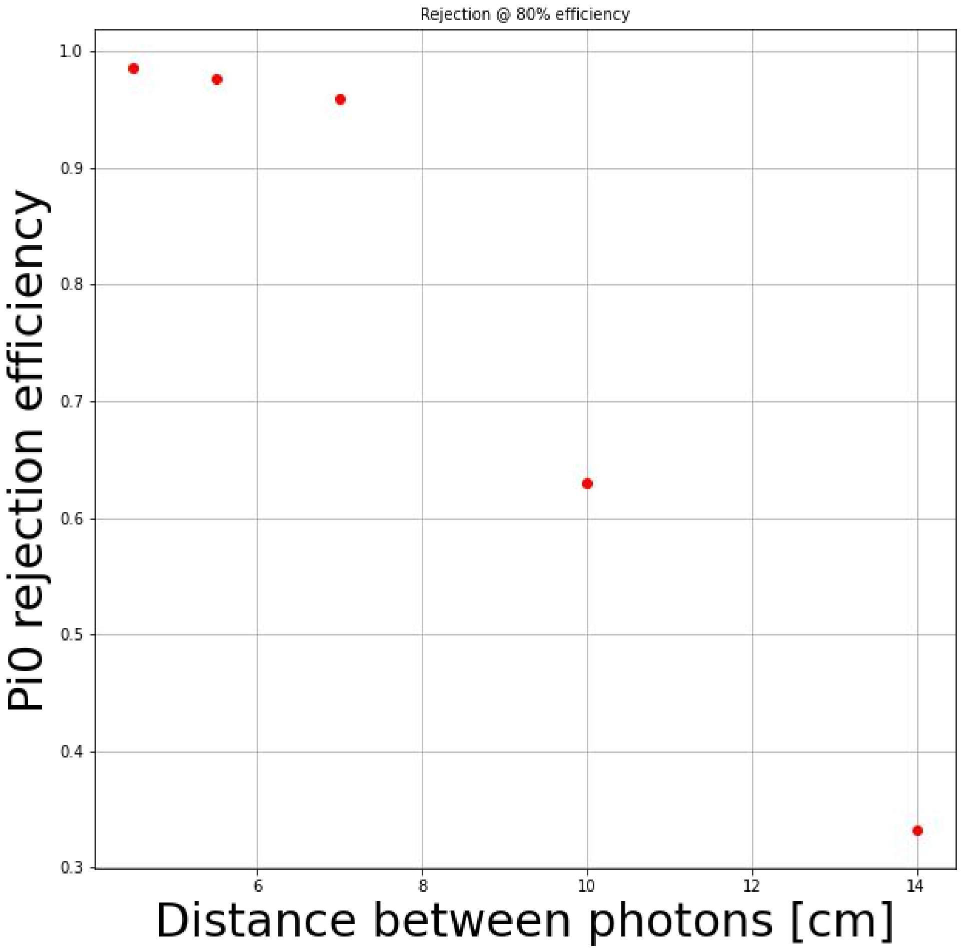
$$\frac{S_1}{S_9} \qquad \frac{S_9 - S_1}{S_{25} - S_1} \qquad \frac{M_2 + S_1}{S_4} \qquad \frac{S_6}{S_9} \qquad \frac{M_2 + S_1}{S_9}$$

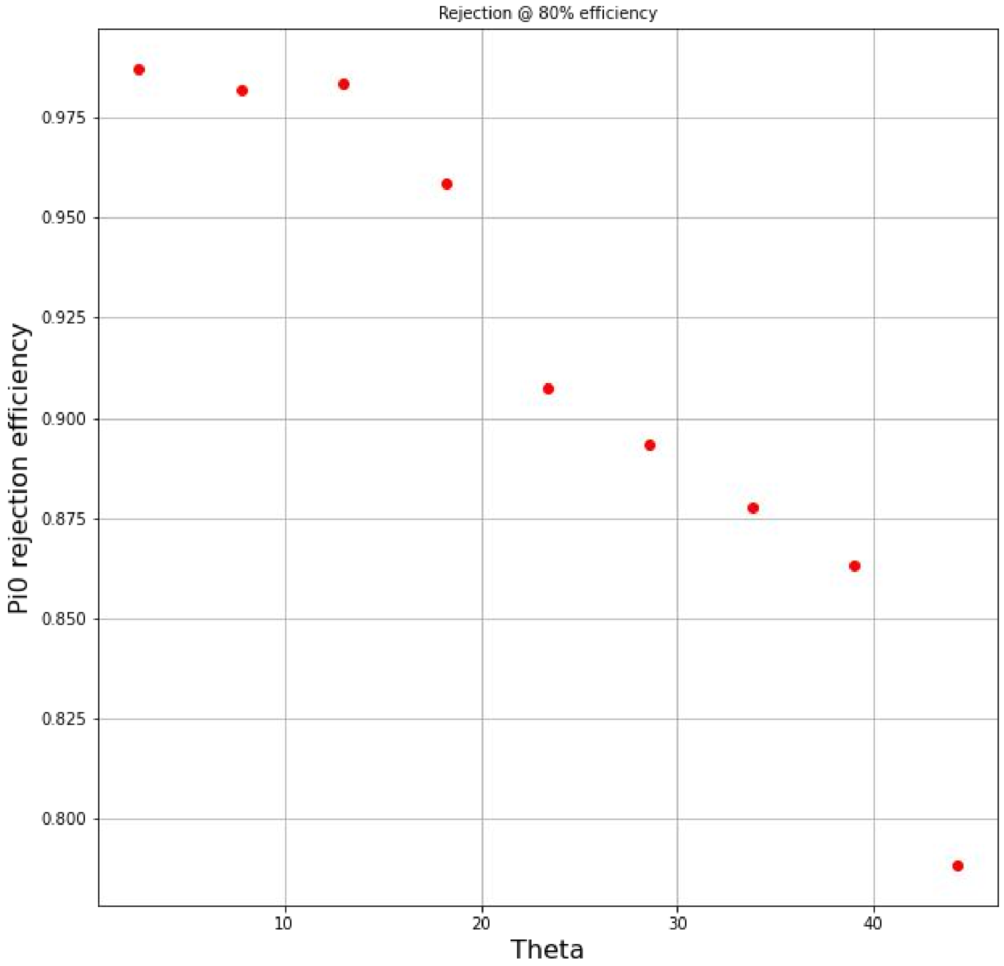
- Angle θ as an input variable (improves separation at high energies)
- Total energy

X,Y ~ θ,φ
S_1, M_2 - 1st and 2nd largest energies
S_9, S_{25} - energy in 3x3, 5x5 region
S_6 - maximum energy in 3x2 region containing S_1 and M_2

14 inputs

Previous results





Test set and train set were parts of the common dataset (2/3: train, 1/3: test)

Effect of energy uncertainty on NN performance

Train dataset: “correct” MC energy, the following taken into account:

- sampling fluctuations (energy deposition in lead/scintillator)
- longitudinal leakage

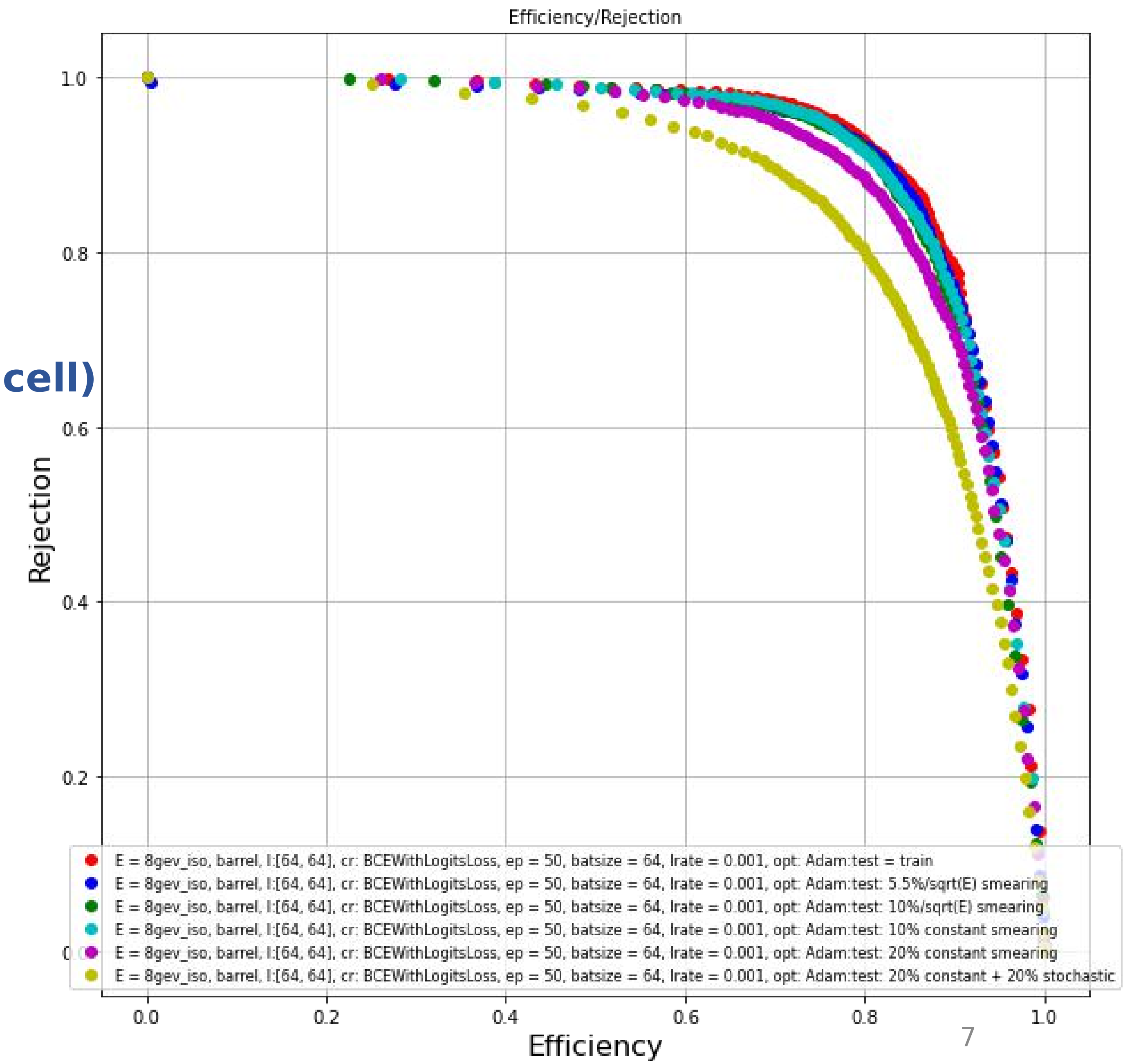
Test dataset: additional smearing (pseudo-real-data):

- miscalibration
- light attenuation
- and other effects not accounted for in MC simulation

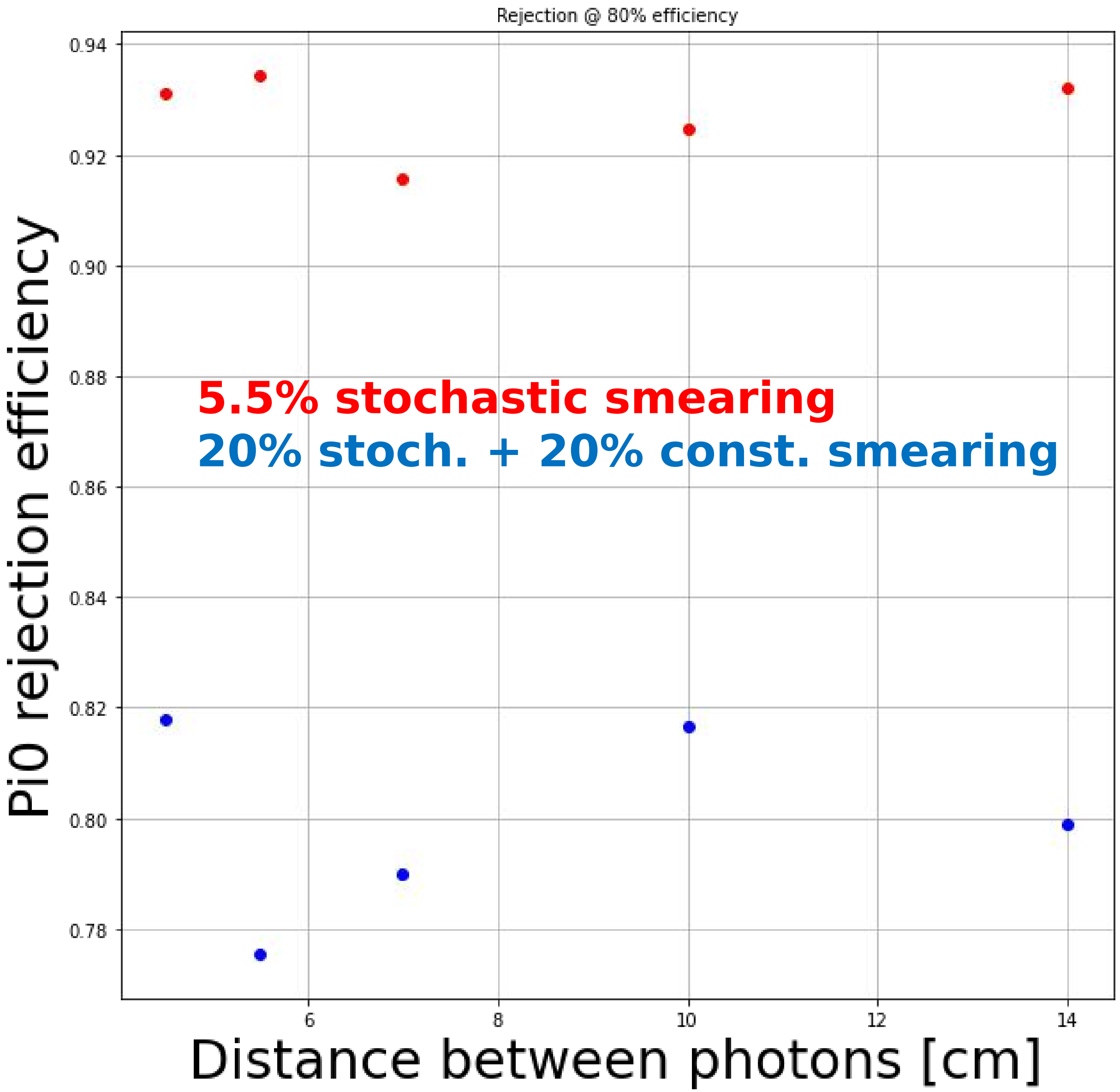
How will the performance change?

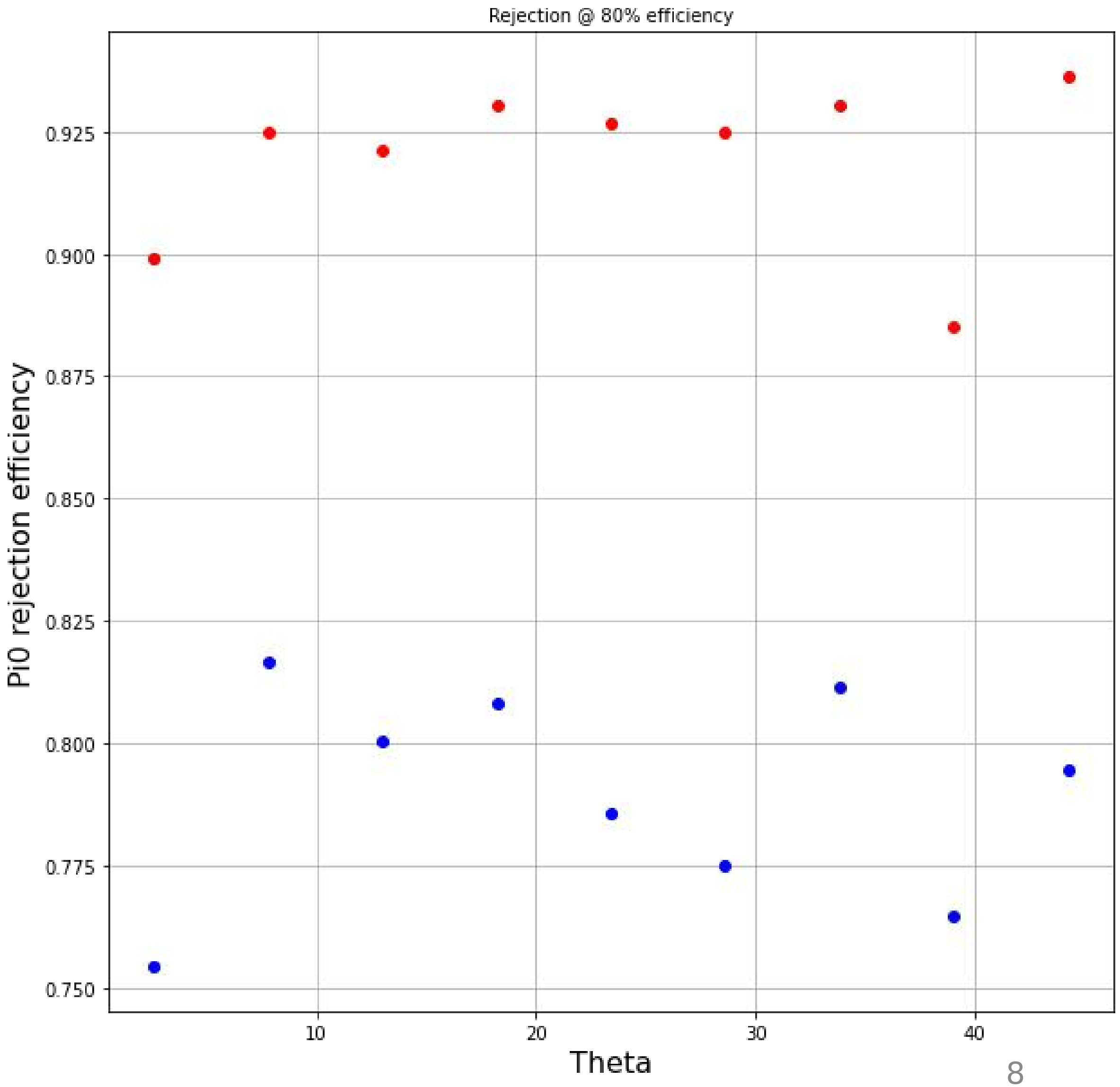
Results

- **test is produced same way as train (but different data points)**
- **5.5%/sqrt(E_{cell}) smearing for test (each cell)**
- **10%/sqrt(E_{cell}) smearing for test**
- **10% smearing for test**
- **20% smearing for test**
- **20% + 20%/sqrt(E_{cell}) smearing for test**



Results


Rejection @ 80% efficiency
Pi0 rejection efficiency
Distance between photons [cm]
0.94
0.92
0.90
0.88
0.86
0.84
0.82
0.80
0.78
6
8
10
12
14
5.5% stochastic smearing
20% stoch. + 20% const. smearing


Rejection @ 80% efficiency
Pi0 rejection efficiency
Theta
0.925
0.900
0.875
0.850
0.825
0.800
0.775
0.750
10
20
30
40

Conclusions

- Additional energy uncertainty of ~10% does not significantly impact the performance of π/γ separation
- Significant impact on π/γ separation performance expected for uncertainty > 15-20%
- After smearing, there is less dependence on incident angle and distance between photons
- Preliminary result: for 5% additional stochastic uncertanity: 92-95% π^0 rejection for 80% γ selection efficiency

Next steps:

- study endcaps, dependence on particle energy and the set of inputs