Energy reconstruction in a liquid argon calorimeter cell using convolutional neural networks

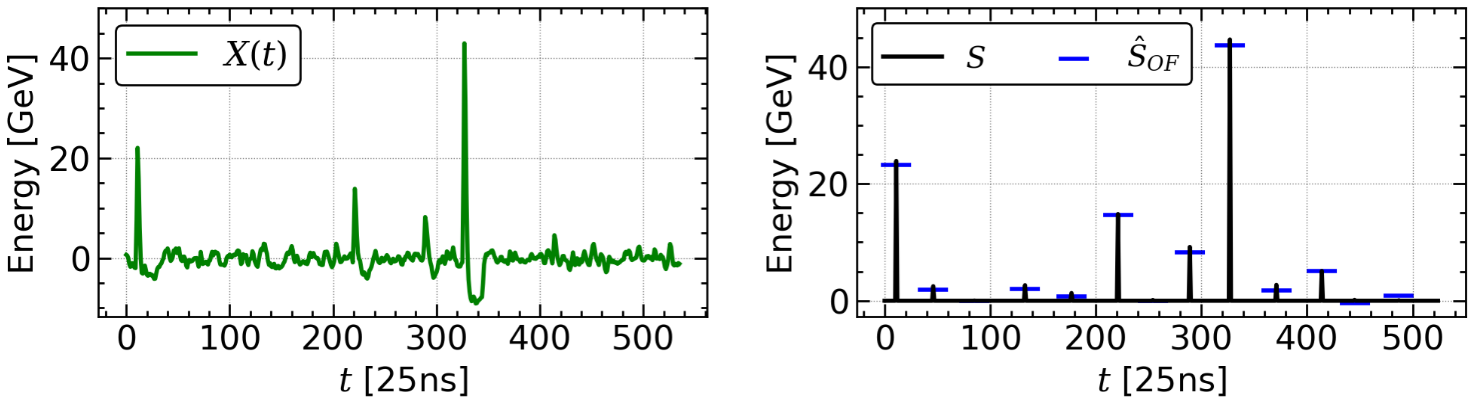
The problem

- The liquid argon ionization current in a sampling calorimeter cell can be analysed to determine the energy of detected particles.
- The inference of energy from measured current in ionization detectors is often a non-trivial process, particularly in the regime of high energy particle physics.
- The conditions in the liquid argon (LAr) calorimeter of the ATLAS detector expected after the high luminosity (HL) upgrade to the Large Hadron Collider (LHC) are such that the precision of energy measurements will suffer a loss in performance.

The goal

- This paper seeks to simulate the conditions of the HL-LHC and the electronic response of a particular detector cell in the hadronic end cap (HEC) subsystem of the ATLAS LAr Calorimeter.
- The presently used algorithm for energy reconstruction from measured signal, known as the optimal filter (OF), will be compared to convolutional neural network (CNN)
- It will be shown that a model architecture trained with an appropriate loss function on simulated data outperforms the optimal filter in relevant metrics.

The total energy E deposited in a HEC detector cell

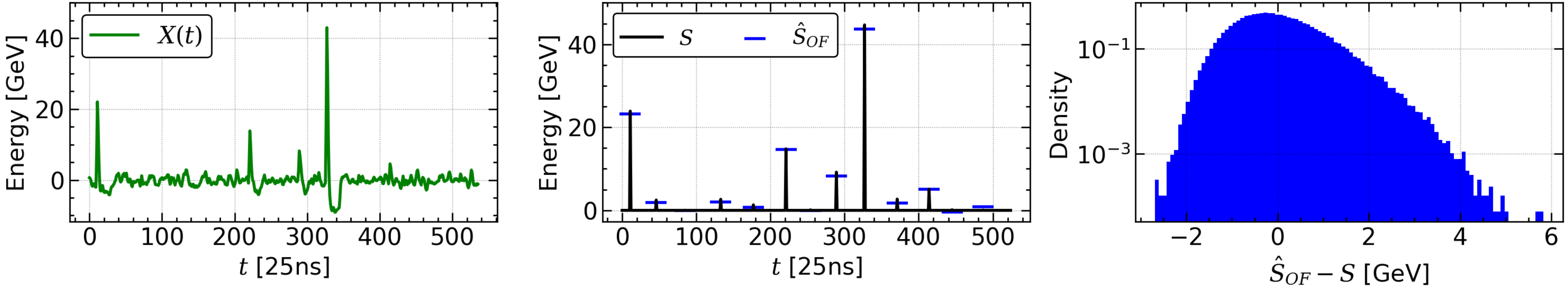


The total energy E deposited in a HEC detector cell at a given BC produces a triangular-shaped ionization drift current pulse which, upon convolution with the electronics chain response, results in the measurable current $X(t)$.

Energy Reconstruction Using the Optimal Filter Technique

- The signal energy is obtained by inferring $S(t)$ from $X(t)$ and is presently obtained using the OF, which produces an estimator given by a weighted sum on $X(t)$ with N coefficients a_i:

$$\hat{S}_{\mathrm{OF}}(t+N) = \sum_{i=1}^{N} a_i X(t+i)$$

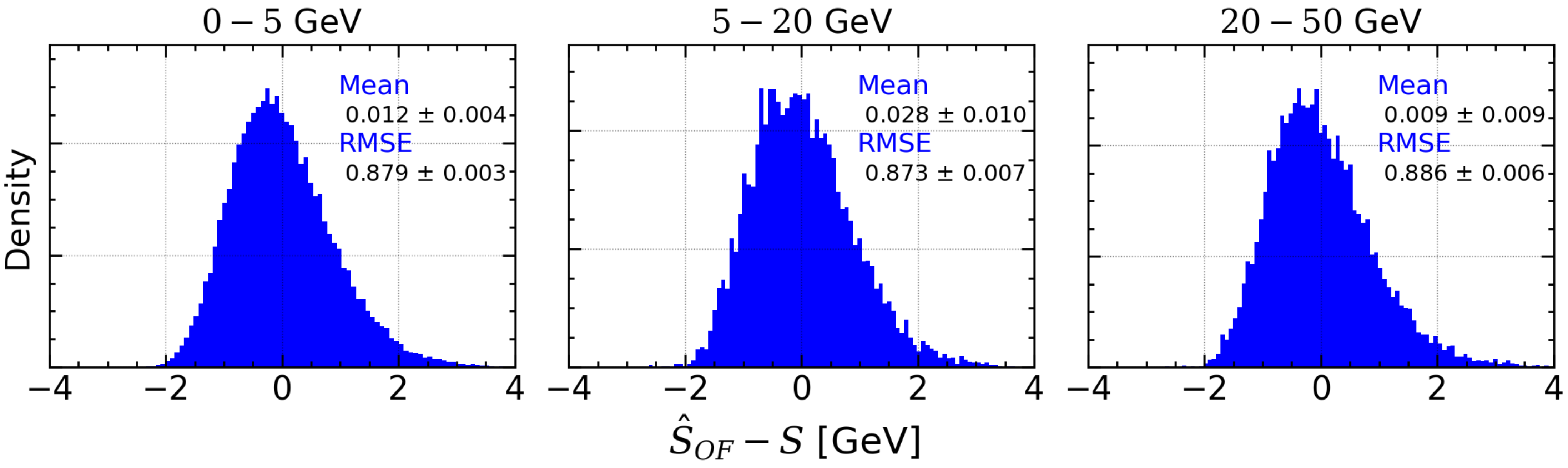


RMSE – Root Mean Square Error

- Root Mean Square Error (RMSE) is the standard deviation of the residuals (prediction errors). Residuals are a measure of how far from the regression line data points are; RMSE is a measure of how spread out these residuals are. In other words, it tells you how concentrated the data is around the line of best fit.

$$RMSE = \sqrt{\overline{(f - o)^2}}$$

- A key feature of the OF is its ability to estimate $S(t)$ with little bias in separate energy regions: this is apparent in he figure, where residuals are separated based on signal magnitude. Such an estimator is said to have no energy range bias;



- The extent to which an estimator makes locally biased predictions can be quantified by the following metric

$$B_L(S, \hat{S}) = \sum_{j=1}^{n} \left| \frac{1}{N_j} \sum_{S \in P_j} \left(\hat{S}(\theta) - S \right) \right|$$

- The P_j's used are given by Table 1.

Subset	P_1	P_2	P_3	P_4	P_5	P_6	P_7
Values [GeV]	0-0.3	0.3-1	1-2	2-5	5-10	10-20	20-50

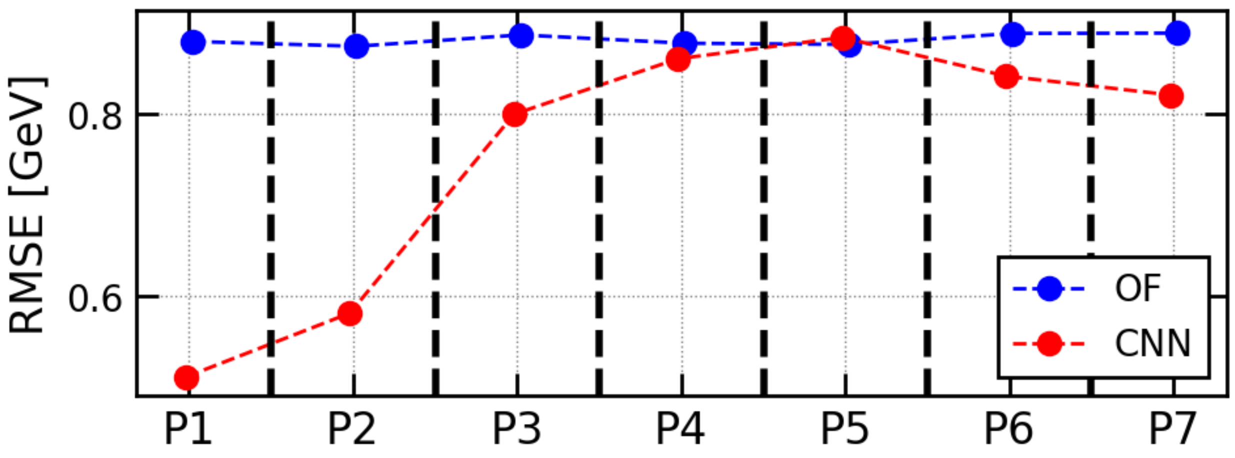
Table 1. Signal partitions.

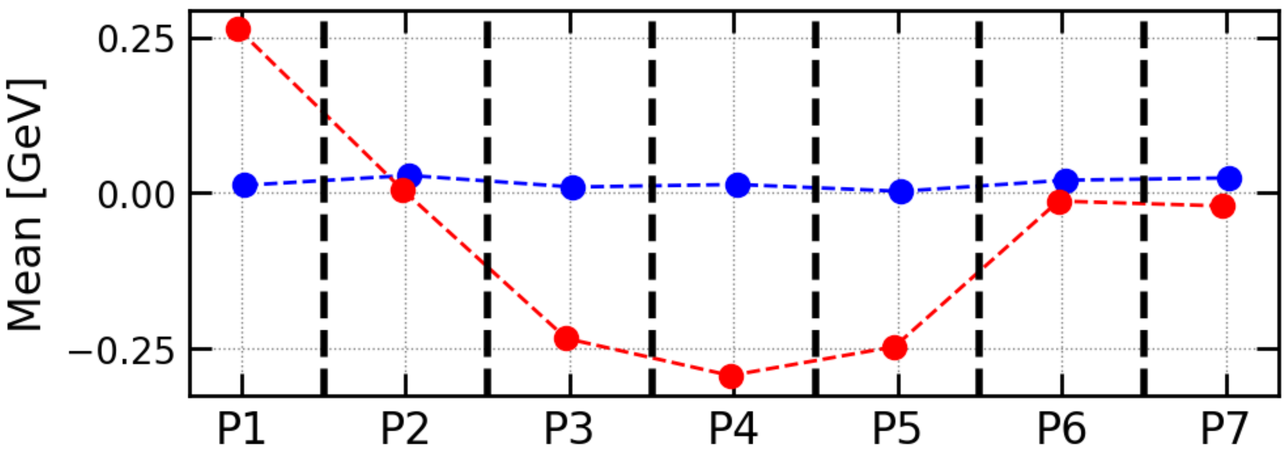
Energy Reconstruction Using a CNN - Using an RMSE Loss Function

Hyperparameters				Results		
Dilations	Filters	Kernel size	Filter Depth	Rank	RMSE [GeV]	B_L [GeV]
(1,1)	3	7	13	1	0.7294±0.0002	1.117±0.015
(1, 1, 2, 2)	2	5	25	2	0.7300±0.0004	1.118±0.007
(1, 1, 1, 2)	2	5	21	3	0.7308±0.0005	1.120±0.009
(1, 2)	3	7	19	4	0.7318±0.0006	1.080±0.020
(1, 1, 2)	2	7	25	5	0.7321±0.0003	1.131±0.005
(1, 3)	3	7	25	6	0.7323±0.0007	1.098±0.017
(1, 2)	3	5	13	7	0.7324±0.0008	1.067±0.014
(1, 3)	3	5	17	8	0.7341±0.0004	1.081±0.006
(1, 1, 3)	2	5	21	9	0.7357±0.0002	1.075±0.006
(1, 3)	4	3	9	11	1.0678±0.0014	2.022±0.017
Optimal Filter				10	0.8810±0.0016	0.108±0.021

Energy Reconstruction Using a CNN - Using an RMSE Loss Function

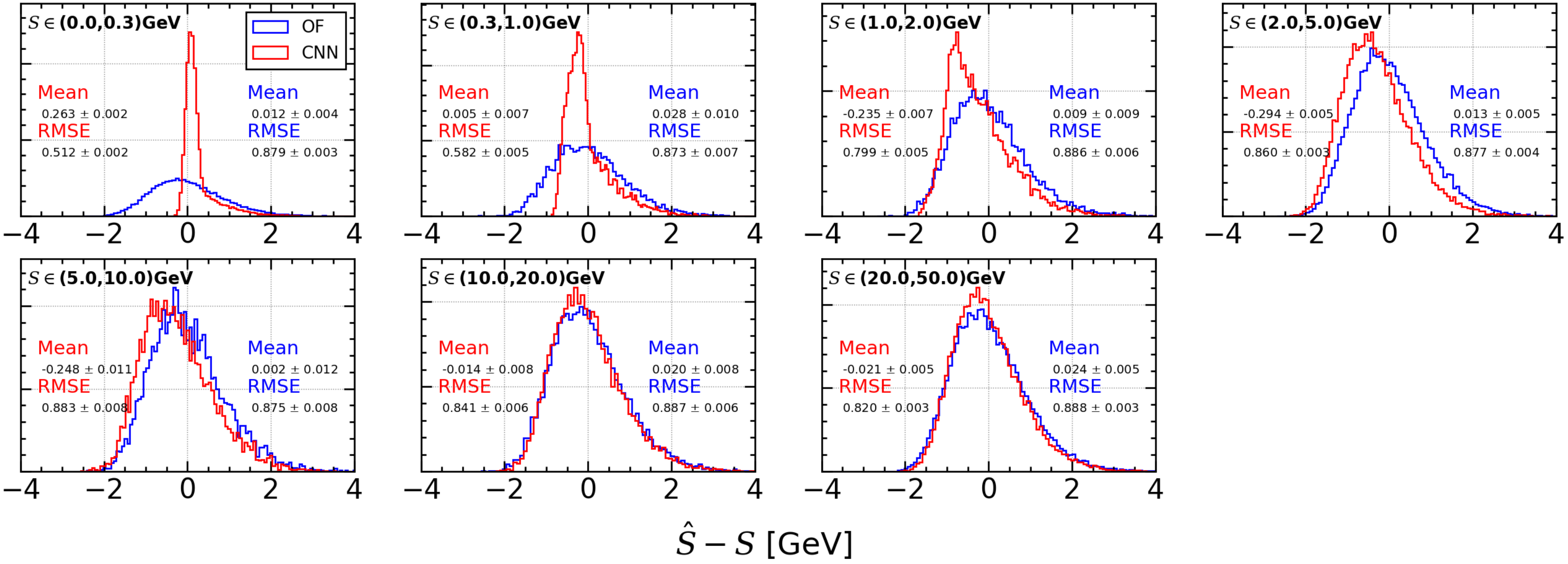
- While the CNNs tend to make predictions with lower RMSE than those of the OF, the large B_L values are indicative of large local biases in CNN predictions.
- This is visible, for example, in the predictions of the CNN model architecture with the lowest RMSE score





Energy Reconstruction Using a CNN - Using an RMSE Loss Function

- Corresponding histograms of residuals in each region P_j are shown
- Compared to the OF, over all reconstructed signals, the RMSE obtained with the CNN is (17.2 ± 0.2)% lower.

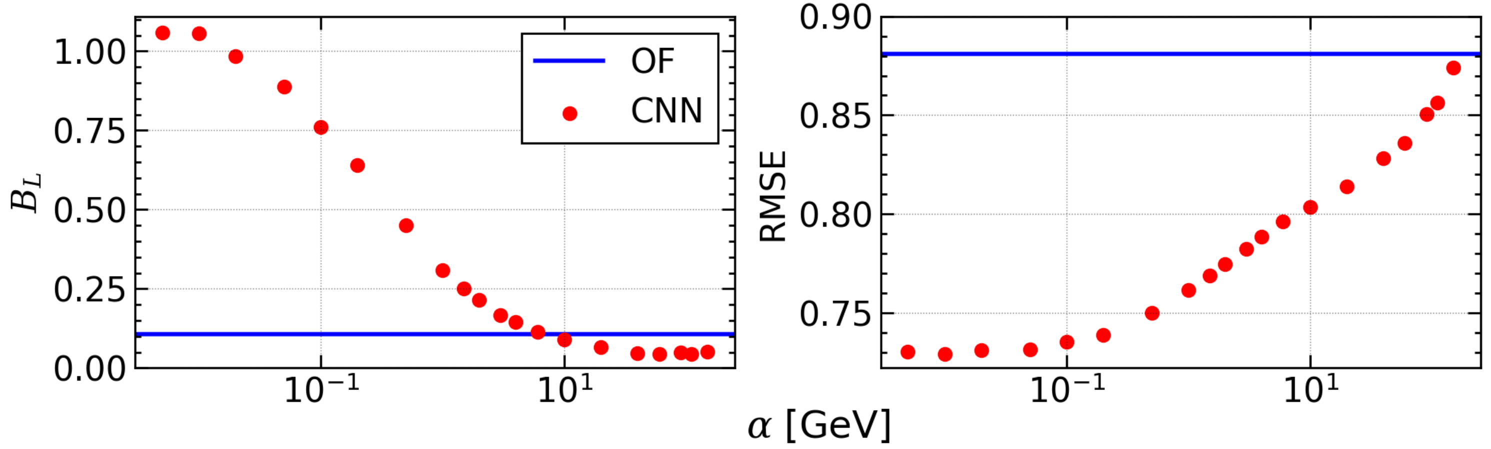


Energy Reconstruction Using a CNN - Using an Alternative Loss Function

- To prevent CNNs from being optimized in such a way that they make locally biased predictions, the following loss function is proposed

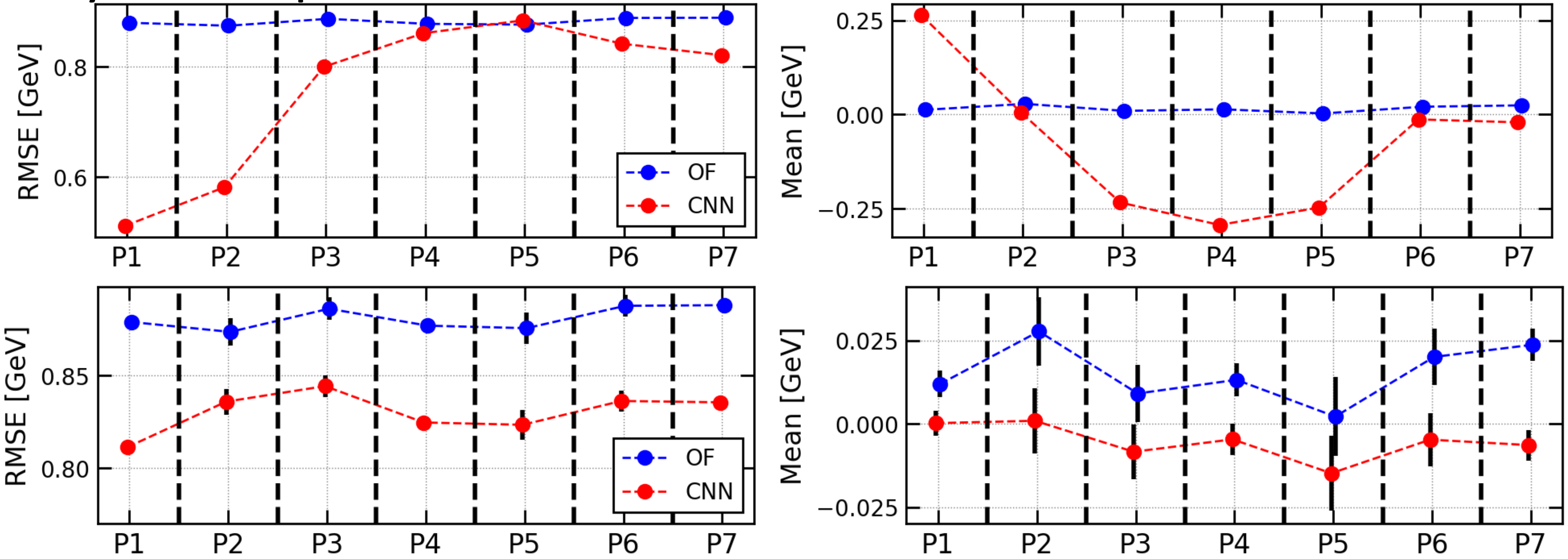
$$L(\theta) = \mathrm{RMSE}(\hat{S}(\theta), S) + \alpha \sum_{j=1}^{n} \left(\frac{1}{N_j} \sum_{S \in P_j} (\hat{S}(\theta) - S) \right)^2$$

- It is shown that a proper adjustment of α results in a CNN predictor with smaller RMSE and B_L than the OF



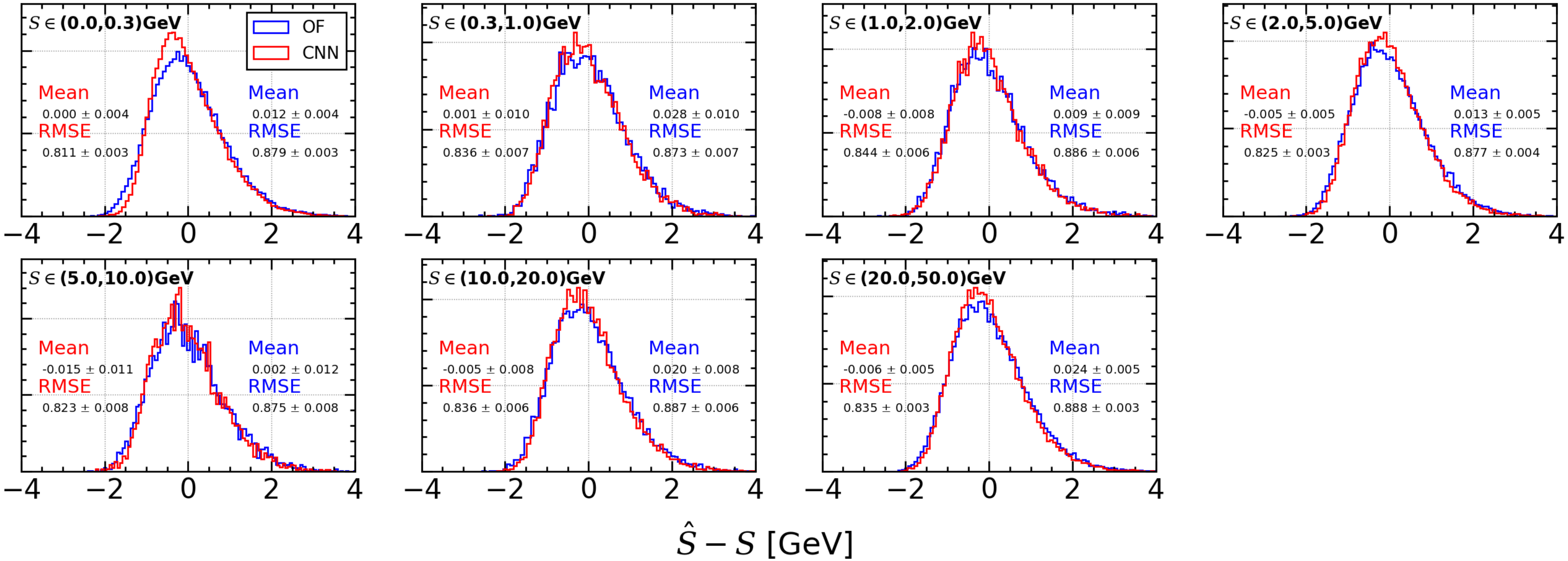
Energy Reconstruction Using a CNN - Using an Alternative Loss Function

- The mean and RMSE of residuals for the top model architecture trained with the value α = 40 are shown
- Compared to the previous loss function it is apparent that the new loss function successfully prevents the CNN model from making locally biased predictions.



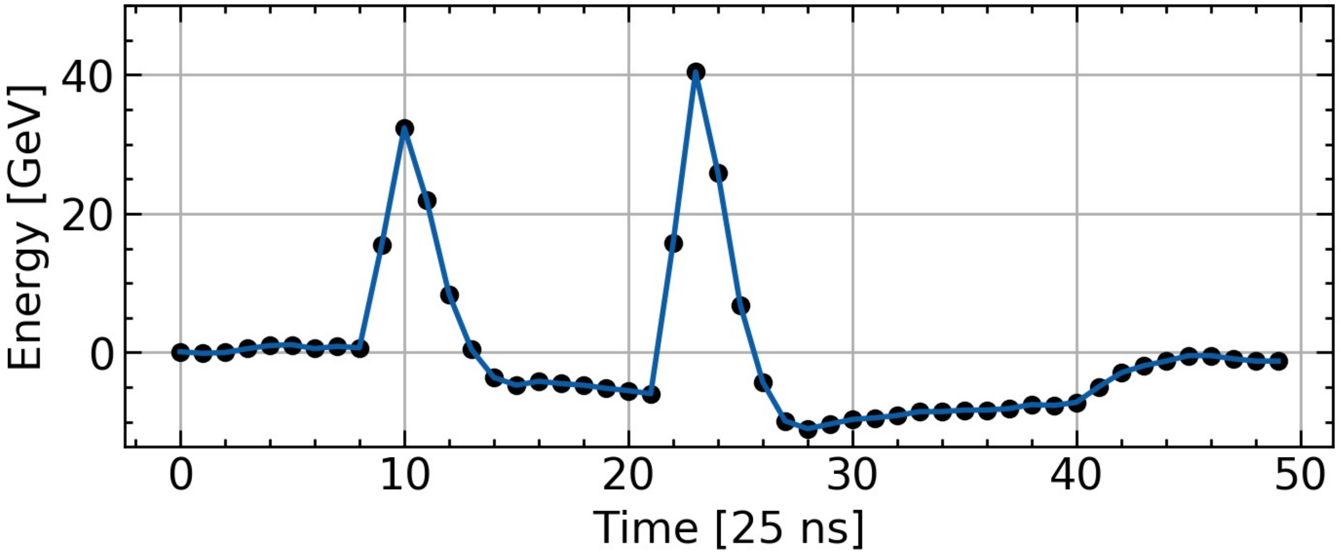
Energy Reconstruction Using a CNN - Using an Alternative Loss Function

- Corresponding histograms of residuals in each region P_j are shown
- Compared to the OF, over all reconstructed signals, the RMSE obtained with the CNN is (6.0 ± 0.2)% lower.



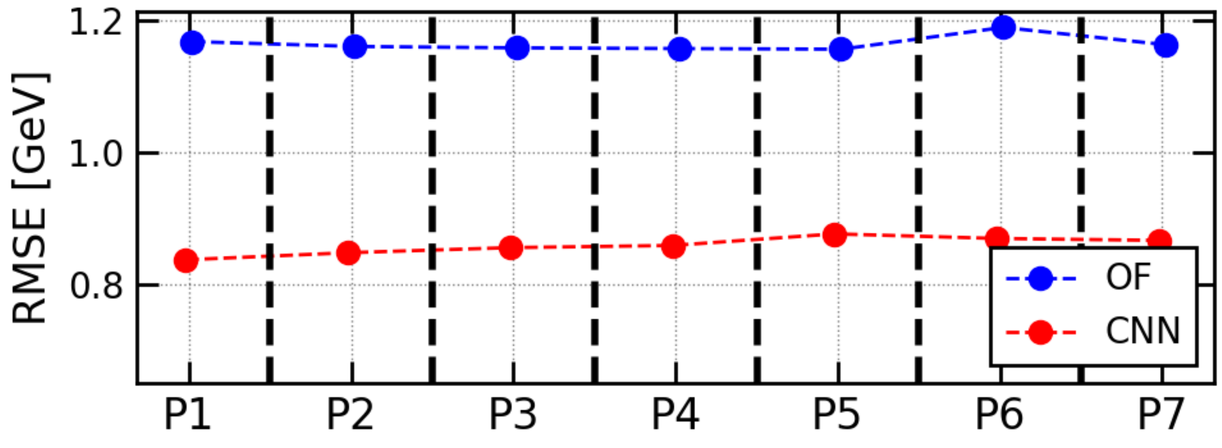
Energy Reconstruction Using a CNN - Overlapping ionization pulses

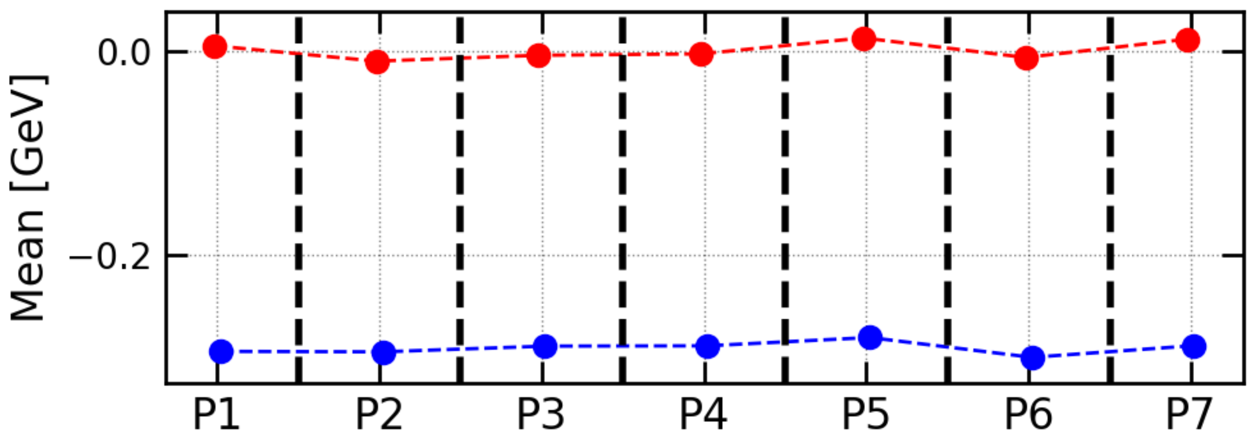
- The optimal hyperparameter configuration and loss function can additionally be used to train a model on a dataset where the signal S is spaced sufficiently close together that overlapping ionization pulses are obtained from consecutive signal events.
- This is simulated by using Γ~U(3BC,50BC); a sample of $X(t)$ obtained using this Γ is shown.



Energy Reconstruction Using a CNN - overlapping ionization pulses

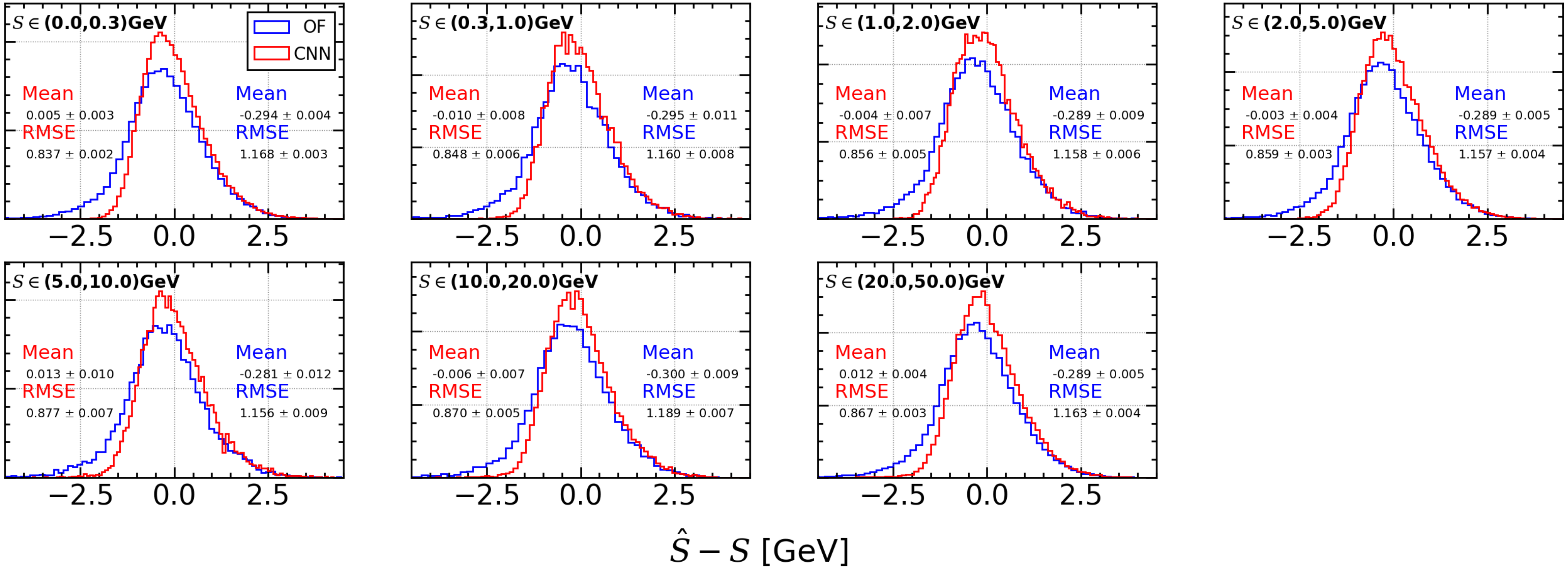
- The performance of the CNN is significantly greater than the OF in this regime.





Energy Reconstruction Using a CNN - overlapping ionization pulses

- Corresponding histograms of residuals in each region P_j:



- The RMSE obtained with the CNN is (26.2 ± 0.2)% lower.

Conclusion

- A standard RMSE loss function was explored as a first option for training CNNs; while the CNN saw a reduction in RMSE of (17.2 ± 0.2)% compared to the OF, the CNN also developed locally biased predictions in different energy ranges.
- A novel loss function for training was developed to address this bias. This loss function eliminated energy range biases and the CNN still outperformed the OF in RMSE by (6.0±0.2)%.
- Signal was then simulated such that consecutive events could lead to overlapping ionization pulses; in this case, the CNN had the greatest performance improvement over the OF, with a (26.2 ± 0.2)% decrease in RMSE and no energy range bias. With the new loss function, a CNN architecture outperforms the OF in RMSE while eliminating energy range biases.