

# DOE Proposal

Dinupa Nawarathne

September 2023

We can employ parametrized classifiers to perform a likelihood ratio test. Let's assume that  $f(x, \theta)$  is a parametrized classifier trained to minimize binary cross-entropy (BCE) loss. Then the likelihood ratio for two parameters  $\theta_0$  and  $\theta_1$ , is given by[1];

$$r(x|\theta_0, \theta_1) = \frac{p(x|\theta_0)}{p(x|\theta_1)} = \frac{f(x|\theta_0, \theta_1)}{1 - f(x|\theta_0, \theta_1)}$$

In this study, we use a deep neural network (DNN) as the classifier. The DNN consists of five hidden linear layers, each containing 64 nodes. The ReLU function is used to activate the hidden layers, along with batch normalization layers. The final output is passed through a Sigmoid activation function. We use reconstructed information, denoted as  $x = \{\text{mass}, p_T, x_p, \phi, \cos \theta\}$ , and particle-level information (generator level), denoted as  $\theta = \{\lambda, \mu, \nu\}$ , as inputs to the classifier. The values of  $\lambda$ ,  $\mu$ , and  $\nu$  are uniformly sampled from the ranges  $[1.5, 2.5]$ ,  $[-0.5, 0.5]$ , and  $[-0.5, 0.5]$ , respectively. The two classes used for classification are given in Table 1.

$x$ values	$\theta$ values	label
mass, $p_T$ , $x_F$ , $\phi$ , $\cos \theta$ generated with $\lambda = 1.$ , $\mu = \nu = 0$ .	$\lambda = [1.5, 2.5]$ , $\mu = \nu = [-0.5, 0.5]$	0
mass, $p_T$ , $x_F$ , $\phi$ , $\cos \theta$ generated with $\lambda = [1.5, 2.5]$ , $\mu = \nu = [-0.5, 0.5]$	$\lambda = [1.5, 2.5]$ , $\mu = \nu = [-0.5, 0.5]$	1

Table 1: Physics variables with labels used in classification.

Note that the  $\theta$  values are same in both classes. But the  $x$  values are generated with different angular coefficients. Training the DNN classifier minimizing BCE loss with Adam optimizer[2] will learn the likelihood ratio for any  $\theta$  parameters. The loss curves are give in Figure 1

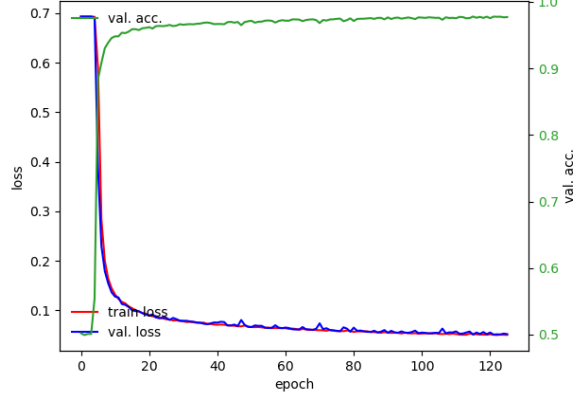


Figure 1: Training loss, validation loss and validation accuracy with epochs.

After training the DNN classifier, we can perform a likelihood ratio test to extract unknown parameters. We can freeze all the weights and biases of the trained DNN classifier and change the parameters  $\theta = \{\lambda, \mu, \nu\}$  by minimizing the BCE loss using the Adam optimizer. The two classes used in the fitting step are given in Table 2.

$x$ values	label
mass, $p_T$ , $x_F$ , $\phi$ , $\cos \theta$ generated with $\lambda = 1.$ , $\mu = \nu = 0.$	0
mass, $p_T$ , $x_F$ , $\phi$ , $\cos \theta$ with unknown $\lambda$ , $\mu$ , $\nu$	1

Table 2: Physics variables with labels used in classification.

The DNN classifier has a minimum loss at the optimal unknown parameter[3]. As an example, the optimal value for  $\mu$  exhibits a minimum loss at the injected value. This is shown in the Figure 2.

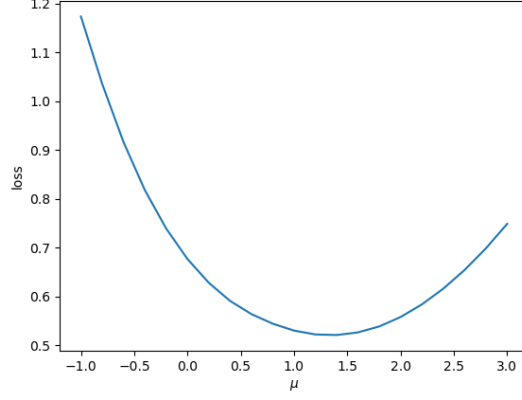


Figure 2: Minimum loss at the optimal value.

Using an gradient decent algorithm we can extract the unknown parameter efficiently. We use Adam optimizer to find the optimal value for unknown parameters. As shown in Figure 3, the fitted parameters converge to the injected values after some epochs.

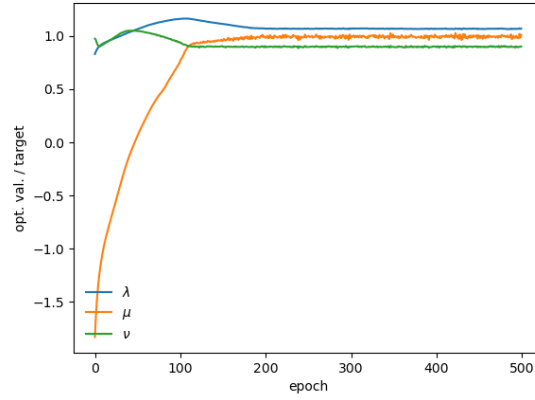


Figure 3: Fitted parameters converge to the injected value after some epochs.

This method allow us to extract the Drell-Yan angular coefficients directly from the reconstructed detector information. Since we can use higher dimensional input features in training, we can use this method to incorporate full phase space when extracting the angular coefficients. This is an advantage

compared to the traditional methods such as RooUnfold.[4]

An early result of this method was presented at the New Perspectives 2023<sup>1</sup> and Neutrino Physics and Machine Learning 2023<sup>2</sup> conferences.

We are investigating how to incorporate generative models to improve the accuracy of this method[5]. We are using conditional variational autoencoders[6] (VAEs) to generate new events with different angular coefficients. The idea is to classify VAE generated events with real events. Since likelihood of these two sets are very close, this method will help us improve the accuracy of the extracted parameters. A simple example of extracting the mean value of the Gaussian distribution is given inFigure 4 and Figure 5.

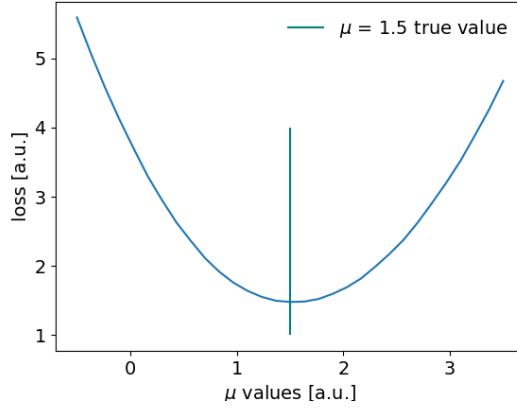


Figure 4: Loss has a minimum at the optimal value using VAE.

---

<sup>1</sup><https://indico.fnal.gov/event/59506/contributions/269963/>

<sup>2</sup><https://indico.slac.stanford.edu/event/8028/timetable/#20230822.detailed>

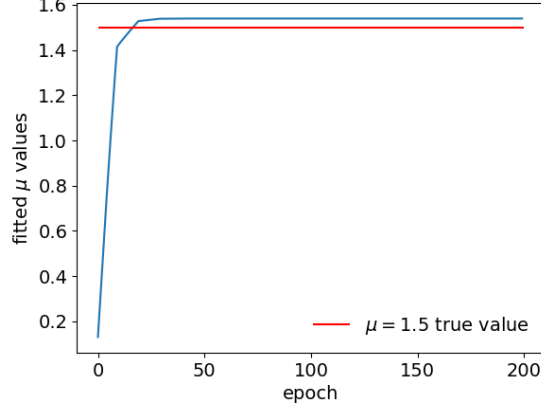


Figure 5: Fitted parameters converge to the injected value after some epochs using VAE.

The code for this study is in GitHub<sup>3</sup> repository.

## References

1. Rizvi, S., Pettee, M. & Nachman, B. Learning Likelihood Ratios with Neural Network Classifiers. *arXiv: 2305.10500 [hep-ph]* (May 2023).
2. Kingma, D. P. & Ba, J. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980* (2014).
3. Andreassen, A. & Nachman, B. Neural Networks for Full Phase-space Reweighting and Parameter Tuning. *Phys. Rev. D* **101**, 091901. *arXiv: 1907.08209 [hep-ph]* (2020).
4. Brenner, L. *et al.* Comparison of unfolding methods using RooFitUnfold. *Int. J. Mod. Phys. A* **35**, 2050145. *arXiv: 1910.14654 [physics.data-an]* (2020).
5. Diefenbacher, S. *et al.* DCTRGAN: Improving the Precision of Generative Models with Reweighting. *JINST* **15**, P11004. *arXiv: 2009.03796 [hep-ph]* (2020).
6. Kingma, D. P. & Welling, M. Auto-encoding variational bayes. *arXiv preprint arXiv:1312.6114* (2013).

---

<sup>3</sup><https://github.com/dinupa1/BMF/tree/main>