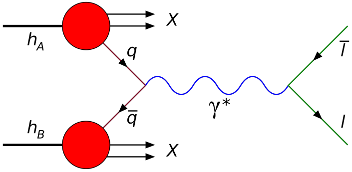
Extraction of Drell-Yan Angular Parameters in *pp* Collisions with a 120 GeV Beam Energy Using a Deep-Learning Unfolding Algorithm

Dinupa Nawarathne

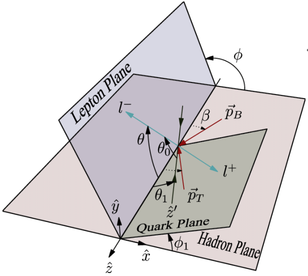
New Mexico State University
Representing the FermiLab SeaQuest/E906 Collaboration
FERMILAB-SLIDES-24-0164-PPD

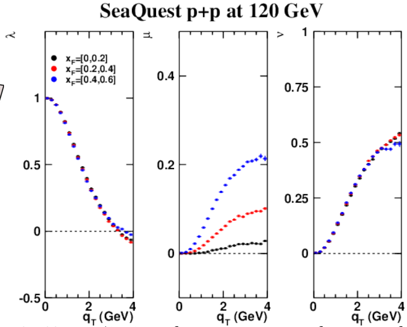
9th July 2024, New Perspectives Meeting 2024

Introduction



Definition of ϕ and $\cos\theta$ angles in the Collins-Soper frame.





NLO pQCD predictions for λ, μ, ν as a function of q_T for SeaQuest experiment.[1]

Drell-Yan angular cross-section;

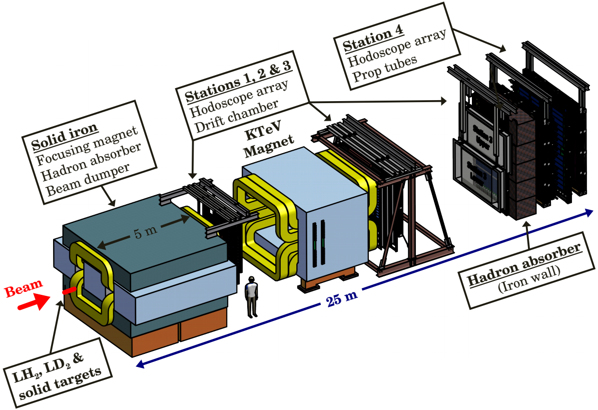
$$\frac{d\sigma}{d\Omega} \propto 1 + \lambda \cos^2\theta + \mu \sin 2\theta \cos\phi + \frac{\nu}{2} \sin^2\theta \cos 2\phi$$

λ, μ, and ν angular parameters, θ and ϕ polar and azimuthal angles in the Collins-Soper frame

[1] W.-C. Chang *et al.*, *Phys. Rev. D* **99**, 014032, arXiv: **1811.03256 (hep-ph)** (2019).

SeaQuest/E906 Experiment

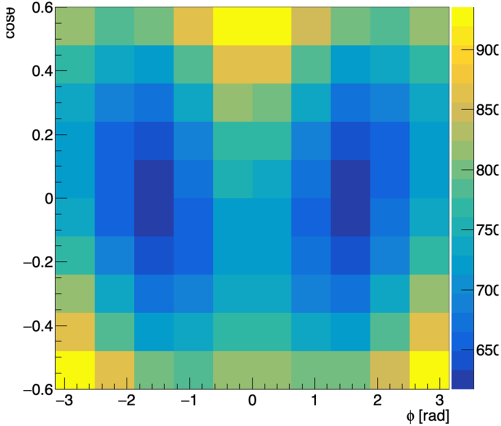
- Fixed target Drell-Yan experiment at Fermilab
- Use 120GeV beam energy from the main injector
- Measure the antiquark structure of the nucleon
- Provides unique access to the vanishing sea quark density at high x_B
- Data collection was concluded in 2017



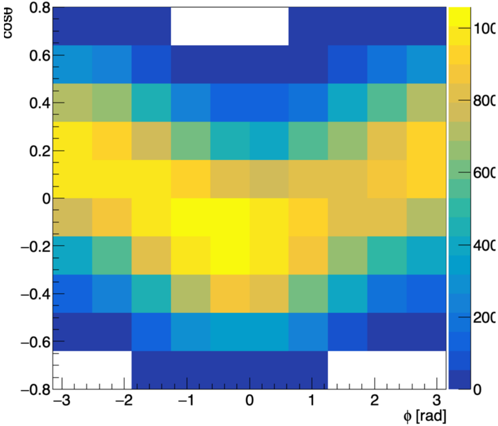
SeaQuest/E906 spectrometer.[2]

[2] C. A. Aidala *et al.*, *Nucl. Instrum. Meth. A* **930**, 49–63, arXiv: 1706.09990 (physics.ins-det) (2019).

Drell-Yan Angular Parameter Extraction



Simulated true level $\cos\theta - \phi$ distributions with $\lambda = 0.8, \mu = 0.1, \nu = 0.2$



Measured $\cos\theta - \phi$ distributions with $\lambda = 0.8$, $\mu = 0.1$, $\nu = 0.2$ using the reconstruction algorithm.

- Reco. $\phi - \cos\theta$ distributions need to be corrected for detector effects (unfolding).

Drell-Yan Angular Parameter Extraction

- Binned unfolding formulation;

$$\vec{m} = R\vec{t}$$

where $\vec{m}$ is the measured distribution (reconstructed), $\vec{t}$ is the true distribution (generated), and R is the response matrix given by:

$$R = Pr(\text{measured i}|\text{truth is j})$$

- In general, R is not invertible. i.e., the unfolding problem does not have a unique solution.
- One common approach to address this problem is iterative Bayesian unfolding (IBU).[3]
- In higher dimensions, these binned methods do not scale well.

[3] G. D'Agostini, *Nucl. Instrum. Meth. A* **362**, 487–498 (1995).

Drell-Yan Angular Parameter Extraction

- We need to build a unbinned model achieving;
  - Background subtraction: Neural Positive Reweighting[4]
  - Corrections for detector smearing/detector efficiencies: OmniFold[5]
  - Parameter extraction using maximum likelihood estimation: RooFit[6]
- We use likelihood ratio estimators to iteratively calculate the weights in the unbinned unfolding method.[7]
- Let s be a classifier (a deep neural network) trained to distinguish samples drawn from $p(x|\beta_0)$ and $p(x|\beta_1)$. Then, the likelihood ratio can be approximated as;[8]

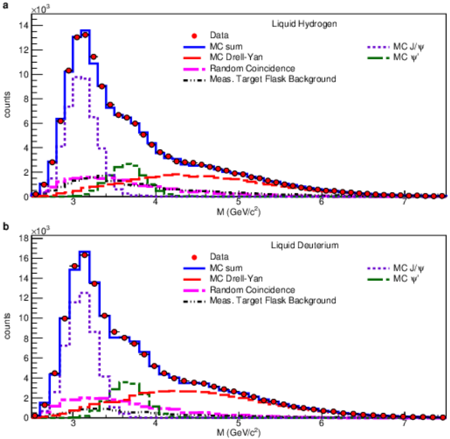
$$\mathcal{L}(x) = \frac{p(x|\beta_0)}{p(x|\beta_1)} \approx \frac{s(x)}{1 - s(x)} \tag{1}$$

[4] B. Nachman, J. Thaler, *Phys. Rev. D* **102**, 076004, arXiv: 2007.11586 (hep-ph) (2020).
[5] A. Andreassen *et al.*, *Phys. Rev. Lett.* **124**, 182001, arXiv: 1911.09107 (hep-ph) (2020).
[6] W. Verkerke, D. P. Kirkby, *eConf* **C0303241**, ed. by L. Lyons, M. Karagoz, MOLT007, arXiv: physics/0306116 (2003).
[7] K. Cranmer *et al.*, arXiv: 1506.02169 (stat.AP) (June 2015).
[8] S. Rizvi *et al.*, *JHEP* **02**, 136, arXiv: 2305.10500 (hep-ph) (2024).

Case Study

- To demonstrate the unbinned unfolding method, we use the SeaQuest/E906 simulation framework to generate Monte Carlo (MC) events. These events are;
  - Generated using the PYTHIA event generator.
  - Generated events were then passed through the SeaQuest/E906 detector simulation (using GEANT4) to obtain the reconstructed detector information.
- Data set features;

Type	λ	μ	ν
Simulation	1.	0.	0.
Pseudo data	0.8	0.1	0.2
Background	0.5	-0.2	-0.1



SeaQuest MC data is in good agreement with real data.[9]

[9] J. Dove *et al.*, *Nature* **590**, [Erratum: Nature 604, E26 (2022)], 561–565, arXiv: **2103.04024 (hep-ph)** (2021).

Model Summary

- Background subtraction;

$$\omega_1 : X^{\text{pseudo data with background}} \rightarrow X^{\text{pseudo data with no background}}$$

- Unfolding;

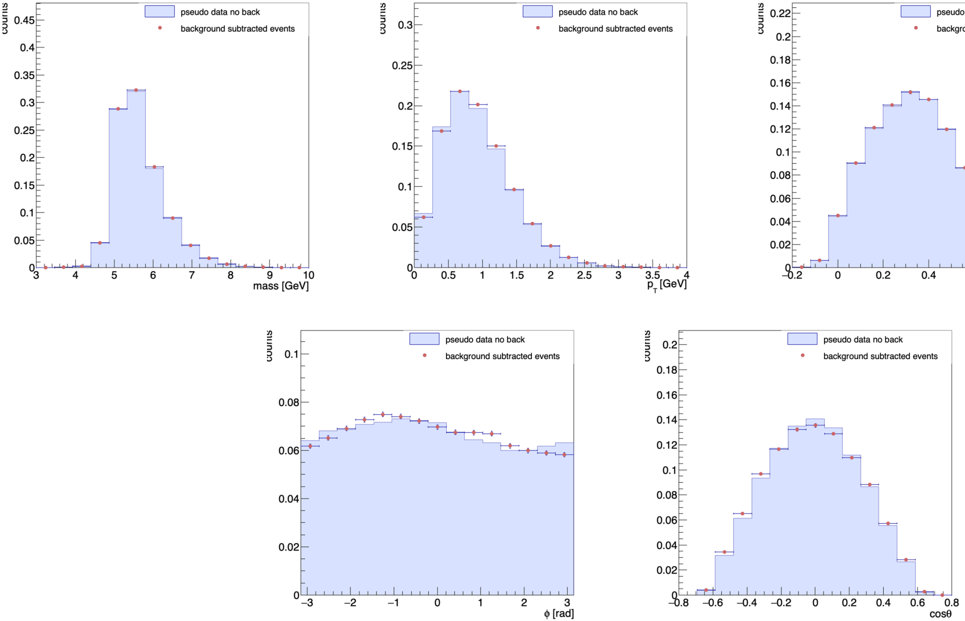
$$\omega_2 : X^{\text{simulation reconstructed}} \rightarrow X^{\text{pseudo data reconstructed}}$$

$$\omega_3 : X^{\text{simulation truth}} \rightarrow X^{\text{pseudo data truth}}$$

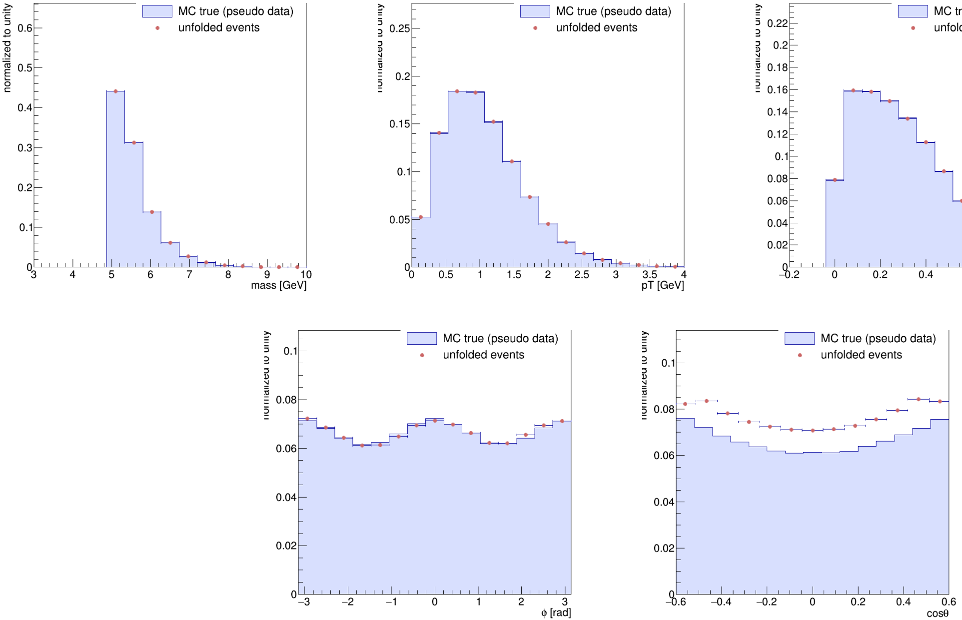
- Parameter extraction;

$$ln(\mathcal{L}) = \sum_i \omega_i ln[1 + \lambda \cos^2\theta + \mu \sin 2\theta \cos\phi + \frac{\nu}{2} \sin^2\theta \cos 2\phi]$$

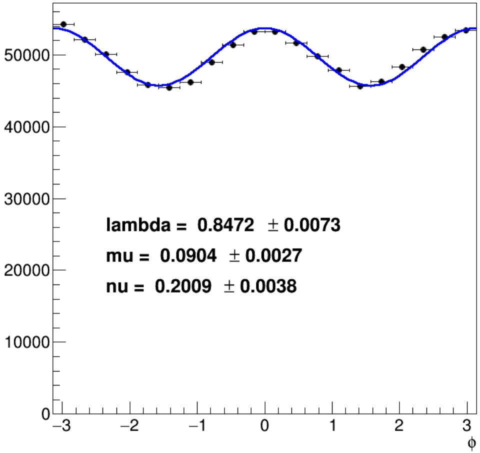
where ω_i are the weights calculated during the unfolding process.

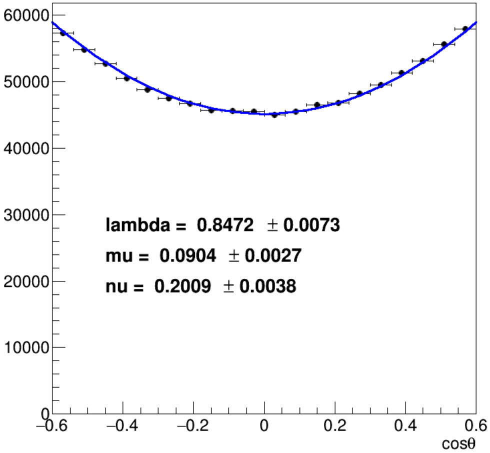


Background subtracted events.



Unfolded events.





Extracted Angular Parameters Using Maximum Likelihood Estimation

Injected values

$\lambda = 0.8, \mu = 0.1, \nu = 0.2$

Neural Network Architecture

- We use feedforward deep neural network for;

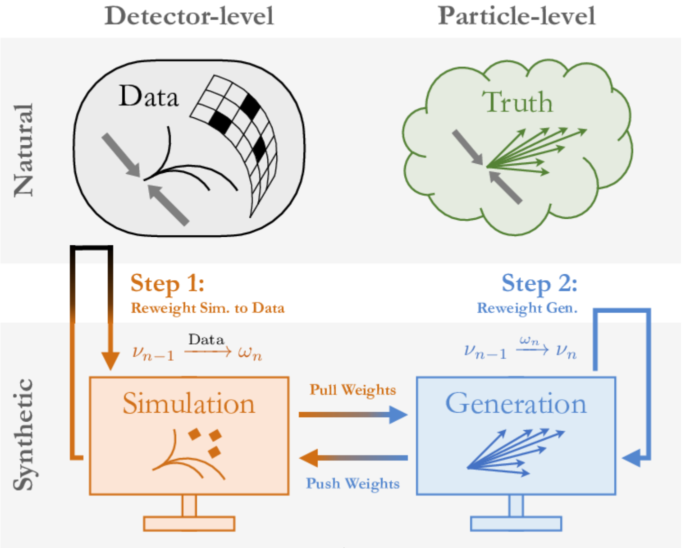
Num. linear layers	5
Num. of nodes per hidden layer	50
Activation function hidden layers	ReLU
Num. of batch normalization layers	5
Num. of dropout layers	2 (p = 0.2)
Final activation layer	Sigmoid
Early stopping patience	10
Learning rate	0.001
Optimizer	Adam
Criterion	weighted BCE
Input features	mass, p_T, x_F, ϕ, $\cos\theta$

- All the models are implemented using the PyTorch framework and trained with NVIDIA A100 GPUs in the Fermilab Elastic Analysis Facility.

Summary

- The Drell-Yan process is an important experimental approach to exploring the partonic structure of nucleons.
- ν provides important information about the intrinsic transverse momentum of partons.
- Deep neural networks are excellent candidates for likelihood estimators due to their feature of approximating complex non-linear functions.
- We can use deep neural networks to unfold the measured (reconstructed) distributions to true distributions.
- Unbinned maximum likelihood estimation can be used to extract the Drell-Yan angular parameters with high accuracy.
- Acknowledgement: This work was supported in part by US DOE grant DE-FG02-94ER40847.

Unfolding: OmniFold



[10] A. Andreassen *et al.*, *Phys. Rev. Lett.* **124**, 182001, arXiv: 1911.09107 (hep-ph) (2020).

Maximum Likelihood Estimation:RooFit

- Consider the log likelihood function;

$$ln(\mathcal{L}) = \sum_i \omega_i ln[1 + \lambda \cos^2\theta + \mu \sin 2\theta \cos\phi + \frac{\nu}{2} \sin^2\theta \cos 2\phi]$$

  where ω_i are the weights calculated during the unfolding process.
- Optimal values for λ, μ and ν can be extracted by minimizing $-ln(\mathcal{L})$.

Background subtraction

- Consider the following event types;

Type	weight sign
Pseudo data + background ($x^{\text{data with back}}$)	+
Background (x^{back})	-

- Typically, the background is combinatorics and it is calculated using the event mixing method.[11]
- We train a classifier ($\hat{s}_{back}(x^{\text{data with back}}, x^{\text{back}})$) to distinguish $x^{\text{data with back}}$, $(y = 1)$ and $[x^{\text{data with back}}, x^{\text{back}}]$, $(y = 0)$ events with the weights. Then weights for the background subtracted events are;

$$W_{x^{\text{data with back.}} \to x^{\text{data no back.}}} = \frac{\hat{s}_{back}(x^{\text{data with back}}, x^{\text{back}}))}{1 - \hat{s}_{back}(x^{\text{data with back}}, x^{\text{back}}))}$$

[11] S. F. Pate *et al.*, arXiv: **2302.04152 (hep-ex)** (Feb. 2023).