

«РЕКОНСТРУКЦИЯ ТРЕКОВ МЕТОДАМИ ГЛУБОКОГО ОБУЧЕНИЯ НА ЭКСПЕРИМЕНТЕ SPD»

Омелянчук С.С
отдел НОВФ, ЛИТ

Научная группа Ососкова Г. А

Филиал МГУ, г. Дубна,
Дубна 2024

Изучение процессов инклюзивного рождения Z -бозона в подходе k_T -факторизации. *M.Nefedov, V.Saleev 2020*

Полученные результаты:

- 1 Полные и дифференциальные сечения рассеяния Z -бозона при энергиях LHC.
- 2 Угловые коэффициенты $A_{0,\dots,4}$ дающие косвенное представление о спиновой структуре адронов.
- 3 Сравнение различных параметризаций партонных функций распределения МКMRW и PB.

Преобрененные навыки:

- Навыки работы в МК генераторах, в частности в KaTie. *A.van Hameren 2018*
- Работа на ОС Linux и написание скриптов формата bash .
- Язык программирования Python и его библиотеки numpy, os, matplotlib.

Публикации:

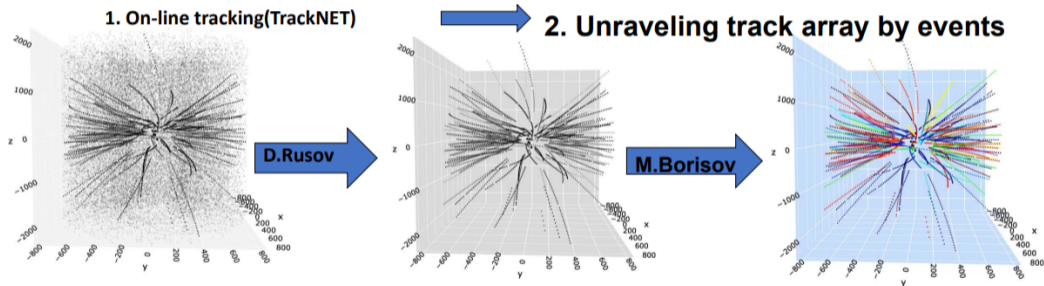
- Spectra of Z Bosons and Lepton Angular Coefficients in the Leading Approximation of the Parton Reggeization Approach with Allowance for the Next-to-Leading-Order Corrections *Phys.Atom.Nucl.* 87 (2024) 3, 237-243
- Z-Boson pT-Spectrum and Lepton Angular Coefficients in the LO High-Energy Factorization with the Real NLO Correction *Phys.Part.Nucl.Lett.* 21 (2024) 4, 695-697

Доклады:

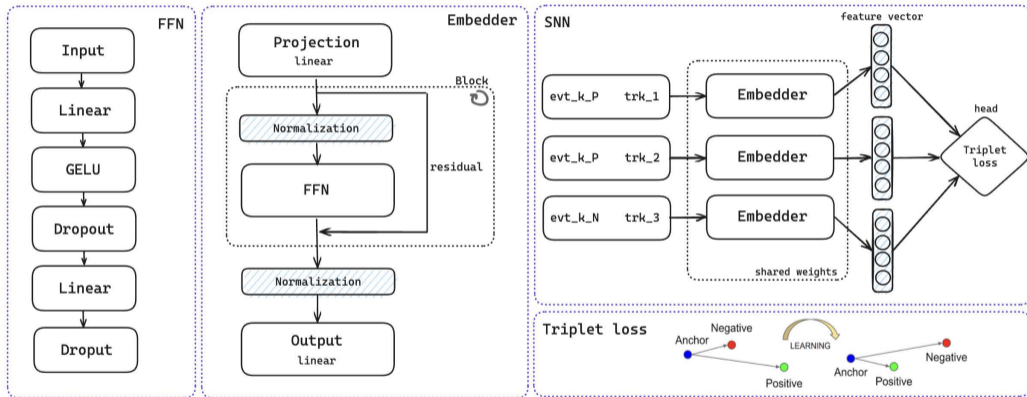
- “Z-boson pT -spectrum and lepton angular coefficients in the LO high-energy factorization with the real NLO corrections”, XIX Workshop on High Energy Spin Physics, DSPIN-23, Dubna, September 4-8, 2023
- “Угловые коэффициенты лептонов в рождении Z-бозонов на БАК в реджевском пределе КХД”, LVI Зимней Школы НИЦ КИ - ПИЯФ, 17 – 22 марта 2024, «Зеленый Бор ЦБ России», Луга
- “Лептонные угловые коэффициенты в рождении Z бозонов в походе факторизации при высоких энергия”, “Научная сессия РАН по отделению ядерной физики”, 1-5 апреля, 2024 год, г. Дубна, ОИЯИ

Задача: классификация трековых событий регистрируемых детектором SPD.

Подход: глубокие нейронные сети для отчистки и классификации.



Процесс обработки экспериментальных данных. Отчистка треков от ложных хитов TrackNETv3, [O.Bakina, D.Baranov... 2022](#) классификация очищенных треков сямской нейронной сетью. [M.Borisova, P.Goncharov... 2024](#)



Архитектура сиамской нейронной сети, использующейся для классификации треков во временном срезе.

■ Внутренние метрики:

$$\text{Silhouette score} = \frac{1}{n} \sum_{i=1}^n \frac{b_i - a_i}{\max(a_i, b_i)} \quad \text{DBI} = \frac{1}{k} \sum_{i=1}^k \max_{i \neq j} \left(\frac{\sigma_i - \sigma_j}{d(c_j, c_i)} \right)$$

где n -количество элементов в данных, a_i -среднее расстояние от точки i до других точек в том же кластере, b_i -минимальное среднее расстояние от точки j до точек в любом другом кластере, k - количество кластеров, σ_i -среднее расстояние от всех точек в кластере i до его центроида, $d(c_j, c_i)$ - расстояние между центроидами кластеров i и j

■ Внешние метрики:

$$\text{Precision} = \frac{TP}{TP + FP} \quad \text{Recall} = \frac{TP}{TP + FN} \quad \text{F1-score} = 2 \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

$$\text{Accuracy} = \frac{TP + TN}{TP + TP + FP + FN}$$

где TP -количество истинно положительных предсказаний, TN -количество истинно отрицательных предсказаний, FP - количество ложно положительных предсказаний, FN - количество ложно отрицательных предсказаний.

■ Алгоритмы кластеризации и сопоставления классов:

KMeans($nclusters=4$), AgglomerativeClustering($Threshold = 2.5$), HungarianMethod

Объяснение результатов: сходимость начинается на 80 эпохах для 40 событий в таймслайсе. *M. Borisova, P. Goncharov... 2024*

Fixed		
Metrics	100 epochs	10 epochs
Silhouette	0,725	0.569
DB	0,655	0.442
Precision	0,811	0.535
Recall	0,843	0.538
F1-score	0,818	0.504
Accuracy	0,895	0.430

UnFixed		
Metrics	100 epochs	10 epochs
Silhouette	0,342	0.578
DB	0,922	0.382
Precision	0,574	0.626
Recall	0,651	0.565
F1-score	0,607	0.559
Accuracy	0,587	0.452

Преобрененные навыки:

- Библиотеки для машинного обучения PyTorch, sklearn, pytorch-lightning.
- Библиотека для работы с данными pandas.
- Библиотека для визуализации данных tensorboard.

!!!Существующая нейронная сеть не подходит для классификации треков с нефиксированным числом событий в временном срезе!!!

Существующая нейронная сеть лучше работает с фиксированным числом событий

Решения проблемы:

- 1 Разработка алгоритма по определению числа вершин взаимодействия с использованием библиотеки [ultralytics](#) и ее пакета [YOLO](#)
- 2 Разработка совершенно новой архитектуры нейронной сети и отказ от старого подхода к обучению

Спасибо за внимание!