Topological Rare Hadron Decay Tagging with DNN: Deep neural net topological tagger for rare hadron decay identification

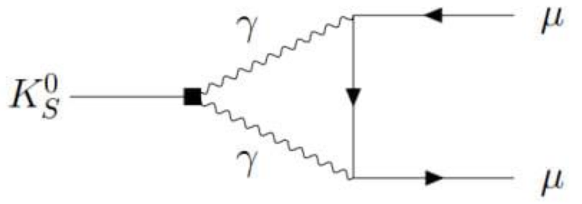
Andrii Len
Taras Shevchenko National University of Kyiv

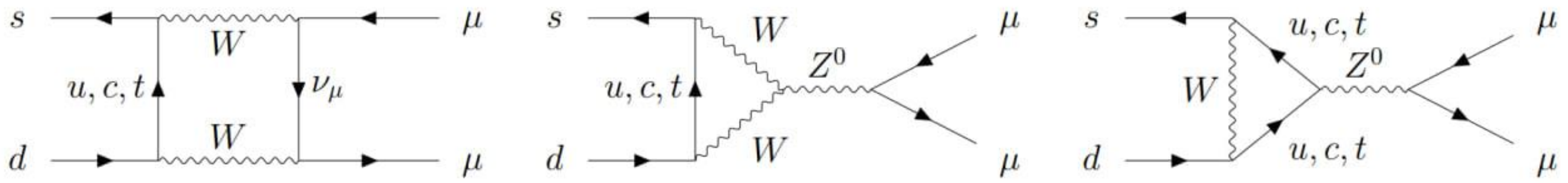
IRIS-HEP Summer Fellowship
Mentors: Dmytro Kovalskyi (MIT, CERN)

July 26, 2024

$K_S^0 \to \mu^+\mu^-$ decay

- The decay $K_S^0 \to \mu^+\mu^-$ is a flavor-changing neutral current (FCNC) process which has not been observed yet.
- In the standard model (SM), this decay is highly suppressed, with an expected branching fraction $B(K_S^0 \to \mu^+\mu^-)_{SM} = (5.18 \pm 1.50_{LD} \pm 0.02_{SD}) \times 10^{-12}$ [G. D'Ambrosio and T. Kitahara, PRL 119 (2017) 201802].
- $K_S^0 \to \mu^+\mu^-$ Decay: A rare decay with the last observed upper limit (B) < 2.4e-10 (LHCb, 95% CL) [LHCb PRL 125 (2020) 231801]) which we want to improve.



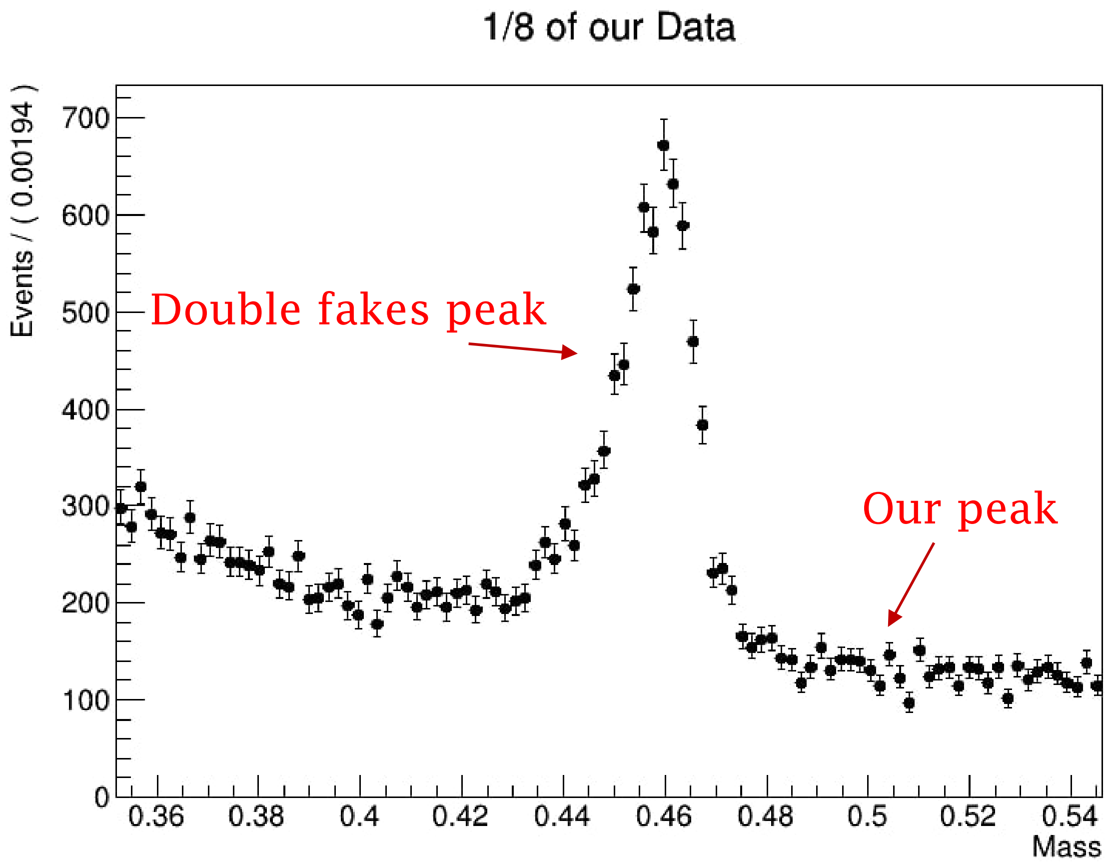


Diagrams representing SM contributions to the K0s → μ+μ− decay amplitude: (top) long-distance contribution, generated by two intermediate photons, and (bottom) short-distance contributions.

Main obstacles

- Rare decay gives a weak signal.

- Existence of other components in the spectrum:
  - Contribution of double fake muons from $K_S^0 \rightarrow \pi^+\pi^- \rightarrow 2\mu 2\nu$ decay (misidentification of pions as muons)
  - Dominated combinatorial background.



Data used

Data used:

- Run3 (2022)
- Run3 (2023)

For signal fit:

- $K_S^0 \to \mu^+\mu^-$ - MC samples
- $K_S^0 \to \pi^+\pi^- \to 2\mu 2\nu$ - MC samples
- Combinatorial Background - fitted on data itself (Parking samples)

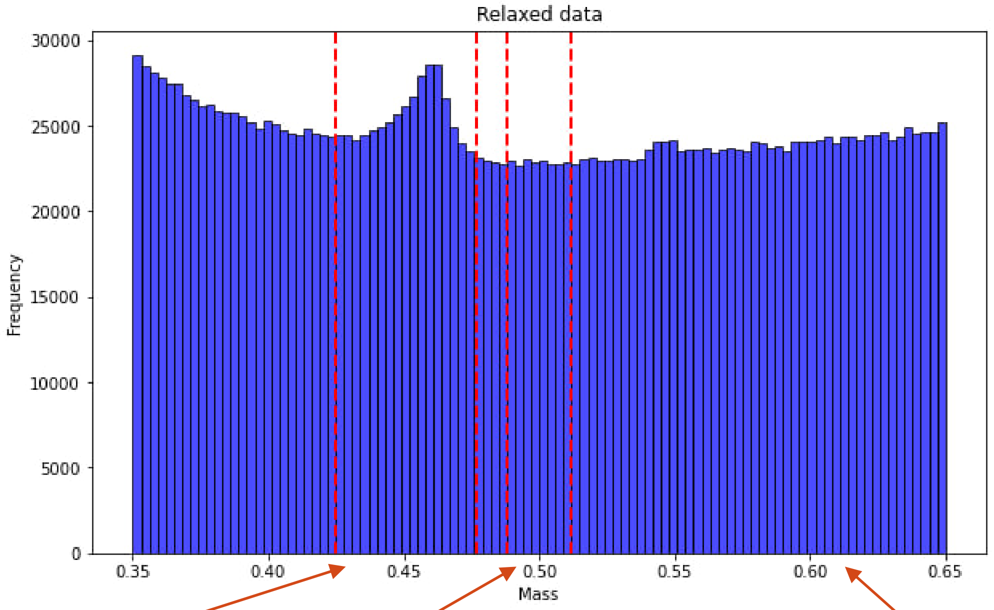
For normalization:

- $K_S^0 \to \pi^+\pi^-$ - ZeroBias

Data selection

- Number of data events: 2 456 790

- With excluded double fake and signal regions - background events left for training: 1 831 235



Double fake region

Signal region

2m2n_left_threshold = 0.425
2m2n_right_threshold = 0.477
mm_left_threshold = 0.488
mm_right_threshold = 0.512

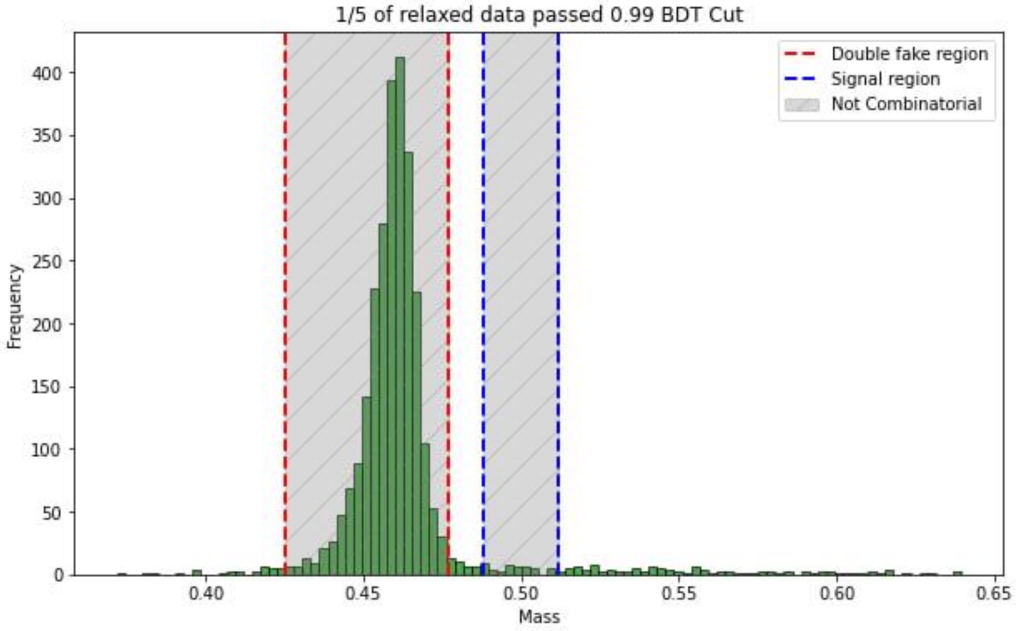
Combinatorial background data

- Total signal MC events: 24 853

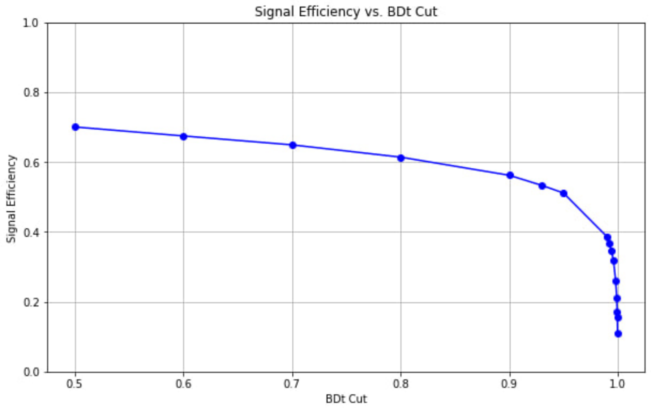
- 0.488 < mass < 0.512 (The region we excluded from data): 22 326

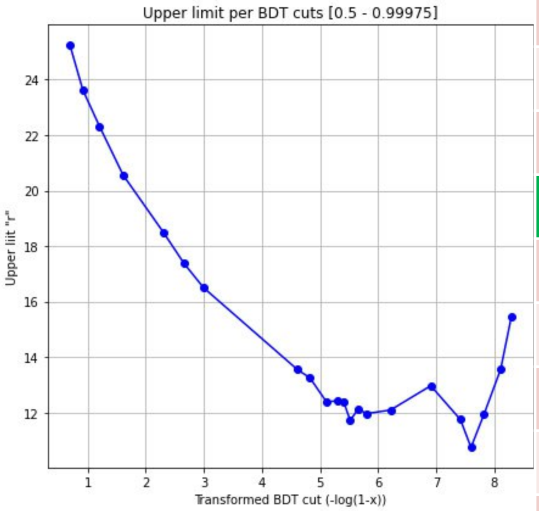
We trained model on 4/5 of data and used 1/5 totally unseen part to make final predictions and new upper limit estimation.

Previous BDT studies



The optimal reliable upper limit numbers are in r < 6 area.





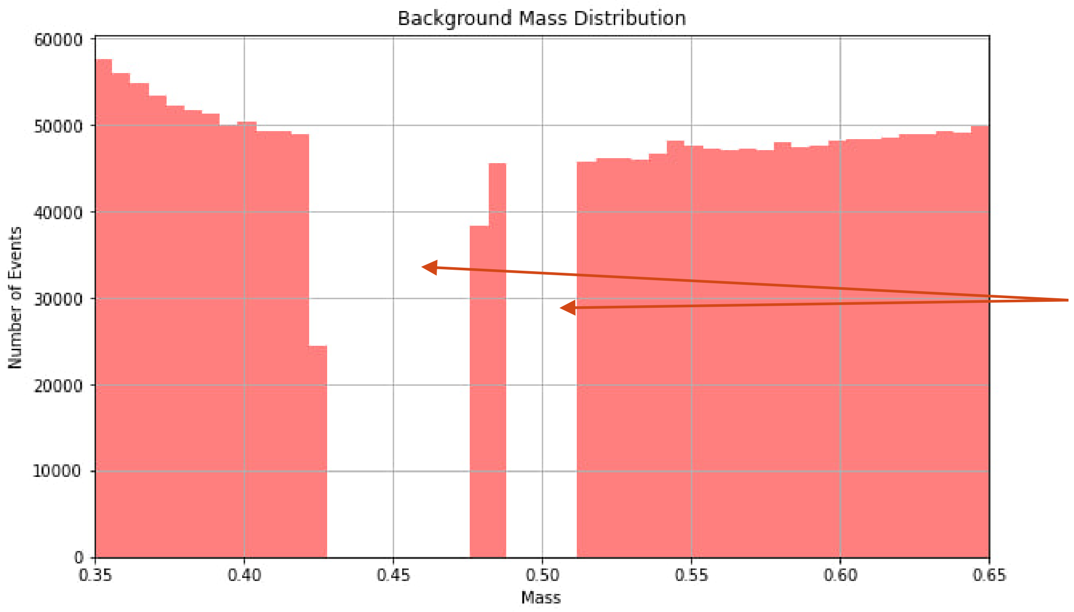
$$\text{Punzi FOM} = \frac{\epsilon}{\frac{n_\sigma}{2}+\sqrt{B}}$$

BDT cut	Expected Signal yield	Expected total (x5) Comb_bkg yield	Punzi FOM	Upper limit
0	4	1830705	0.001	
0.5	2.8	2205	0.015	12.5
0.6	2.7	1755	0.016	11.7
0.7	2.6	1430	0.017	11
0.8	2.45	1115	0.019	10.1
0.9	2.24	670	0.022	9.1
0.93	2.13	555	0.022	8.5
0.95	2.04	465	0.023	8.1
0.99	1.54	170	0.028	6.6
0.992	1.46	140	0.029	6.4
0.994	1.38	105	0.031	6
0.996	1.27	65	0.035	5.5
0.998	1.04	45	0.034	5.8
0.999	0.84	30	0.032	6
0.9994	0.68	15	0.034	5.8
0.9995	0.62	5	0.044	5.3
0.99975	0.44	5	0.031	7.6

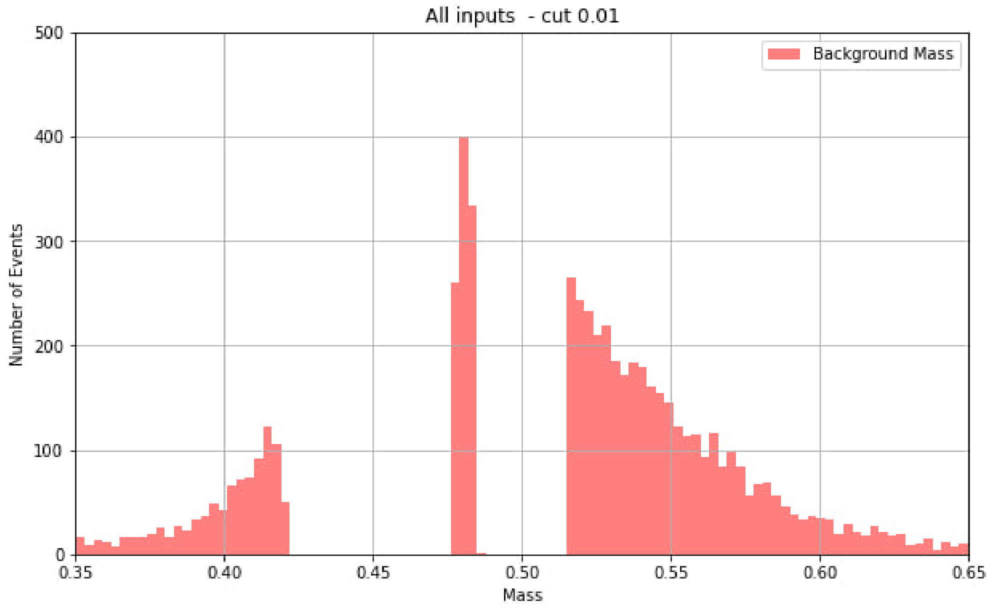
How to improve sensitivity?

- ✓ Expected Upper limit with BDT improved from "r" < 250 to about "r" < 6 (BF $< 1.32 * 10^{-9}$).
- ❖ Though it is not good enough, the upper limit without combinatorial background is at the acceptable level, i.e. we need a better way to reject combinatorial background.
- ❖ We will also try to improve sensitivity of our analysis by increasing signal efficiency.
- ❖ We have several ways:
  - Look into DNNs to find a better way to reject Combinatorial Background.
    - ! *Try different architectures including regular fully connected Networks, Graph Neural Networks, etc.*
    - ! *Try not only Data vs MC training but also Data vs Data.*
  - Explore relaxing kinematic and muon id cuts.
  - More data till the end of the year (2024 Data).

To be careful

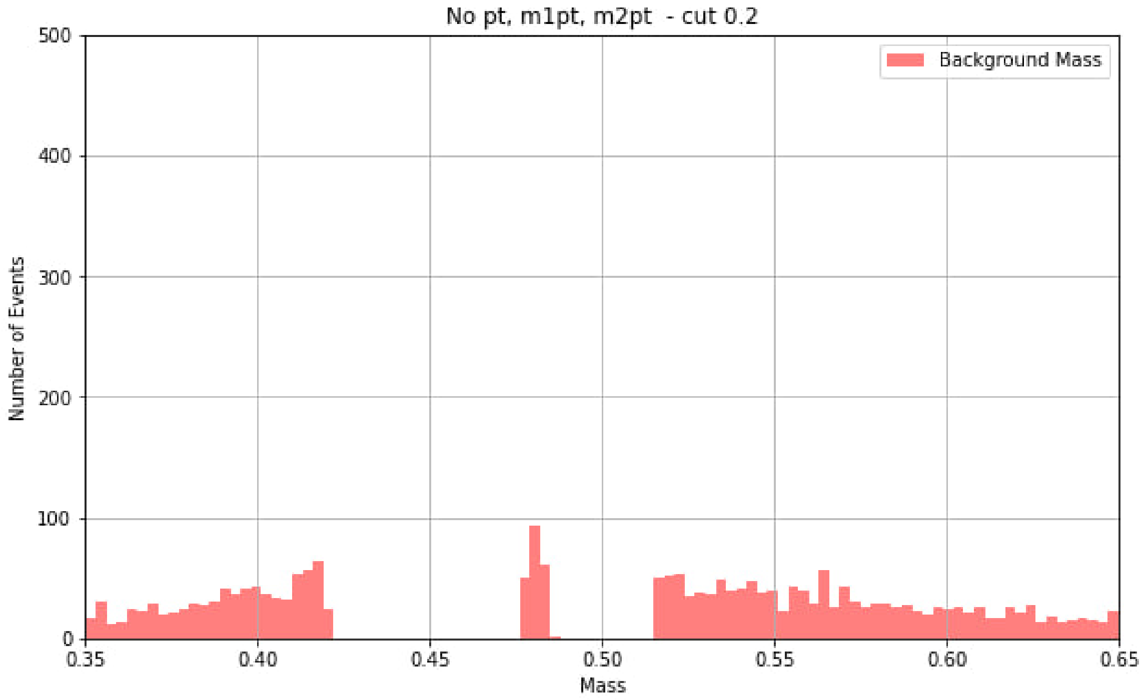


Excluded doublefakes and signal regions from data to obtain combinatorial background events.



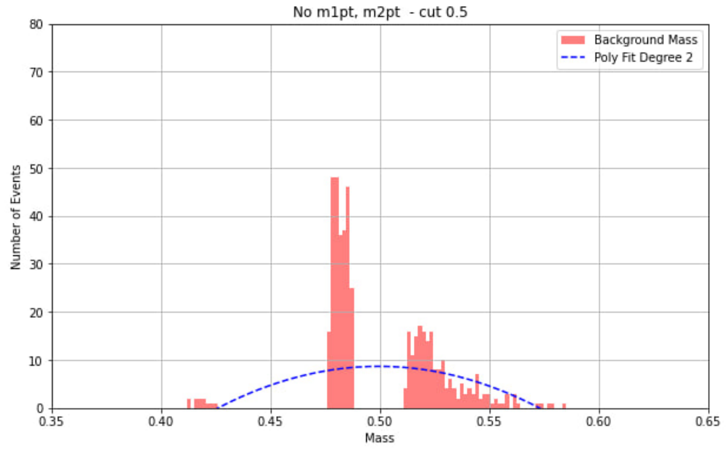
All inputs - clear dependence on mass after applying DNN.

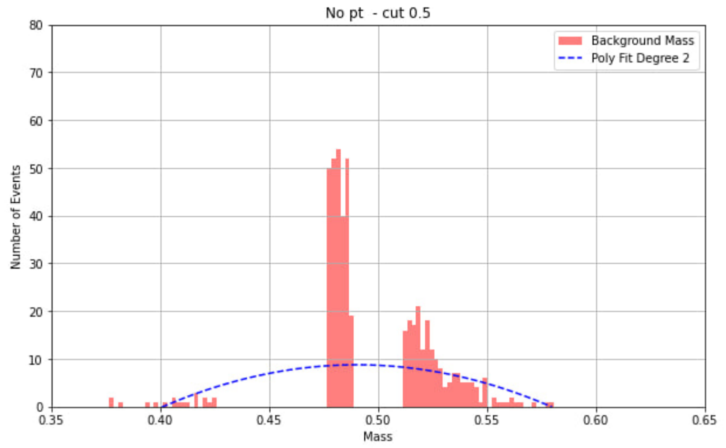
Removed pt, m1pt and m2pt



Almost no dependence on mass

One by one this selections don’t work





- Interesting fact is we can see this only by removing *pt*, *m1pt* and *m2pt* at the same time. Removing just *m1pt + m2pt* or just pt does not have effect.

Summary and goals

Summary:

- The last $K_S^0 \to \mu^+\mu^-$ upper limit was published by LHCb. We want to improve it.
- Using data from Run3.
- Several studies were already performed which increased sensitivity though not enough.

Goals:

- Learn new algorithms and Deep Neural Network architectures.
- Perform robust studies on topological information of our decay to get the most informative features.
- Develop advanced models to surpass previous accuracy and by that improve sensitivity of our analysis.

Thank you for attention!