Transformer based classifier of $B_s^0 \to \tau^+\tau^-$ decays in $Z \to b\bar{b}$ events at FCC-ee

DPG Erlagen 2026
March 19 2026

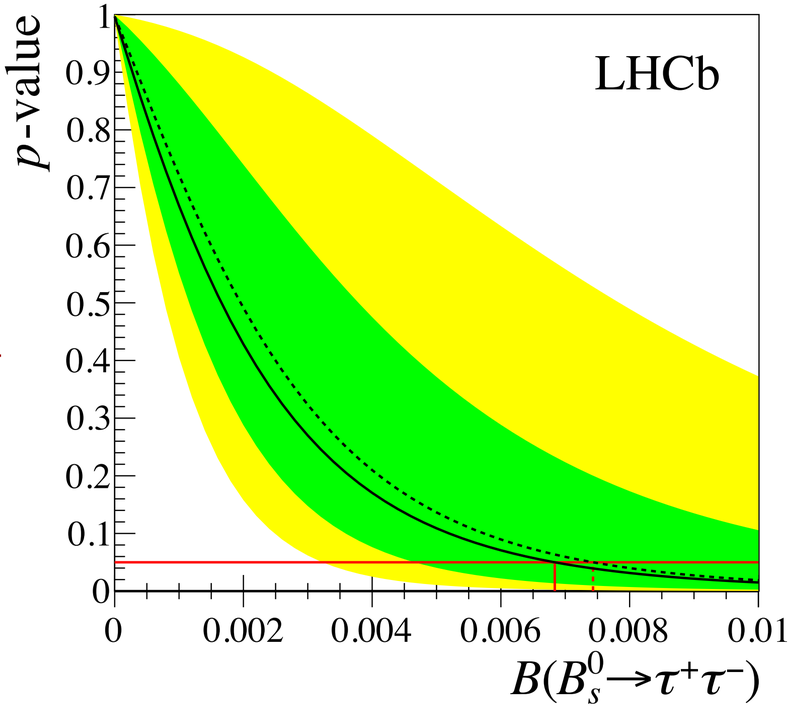
Alejandro Q. Triviño, *Anuar Sifuentes, Jan Kieseler, Markus Klute, Joscha Knolle, Samuel Wyrowski and Xunwu Zuo*

Motivation of study

Experimental limit

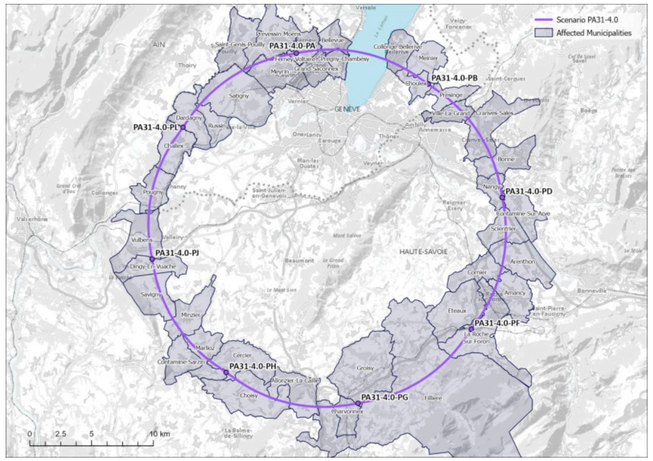
- $B_s^0 \to \tau^-\tau^+$ **rare decay** in the SM
  - $\mathcal{B}(B_s^0 \to \tau^-\tau^+) = (7.73 \pm 0.49) \times 10^{-7}$ ← SM prediction
- Best experimental limit:
  - Run 1 LHCb experiment $\mathcal{B}(B_s^0 \to \tau^-\tau^+) < 6.8 \times 10^{-3}$ at 95% CL
  - HL-LHC Upgrade LHCb expected to be $\mathcal{B}(B_s^0 \to \tau^-\tau^+) <$ 2.6 to 5×10^{-4}
- The measurement of the $\mathcal{B}(B_s^0 \to \tau^-\tau^+)$ is out of range of all current experiments. **What about FCC-ee?**
- An important measurement in the flavour sector, as previous results have shown tensions with the principle of l**epton flavour universality.**

arXiv:1703.02508v2

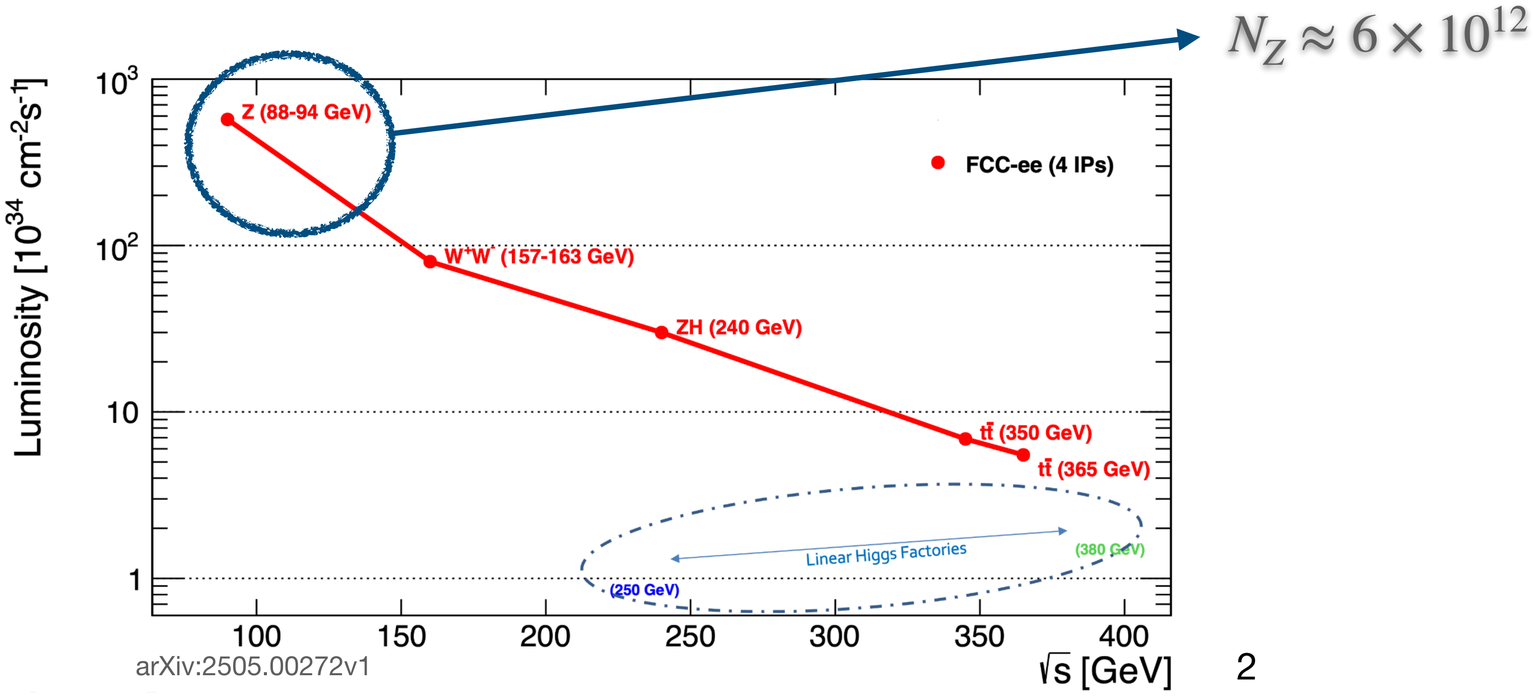


The Future Circular Collider (FCC-ee)

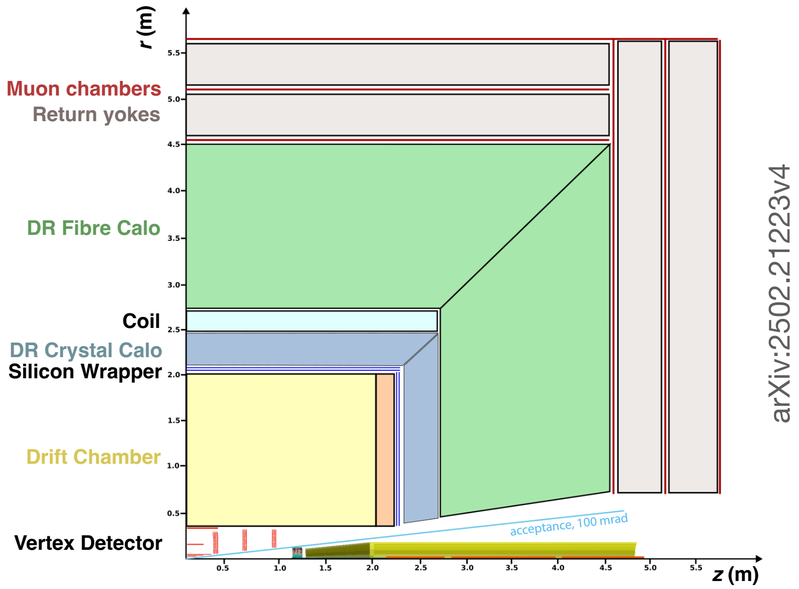
- Electron positron collider proposed to build at CERN
- Feasibility studies: expected approval by CERN Council ~2028 and operation ~ 2045 to 2048.
- Innovative Detector for Electron–positron Accelerators (IDEA) detector concept



90.7 Km of circumference



arXiv:2505.00272v1

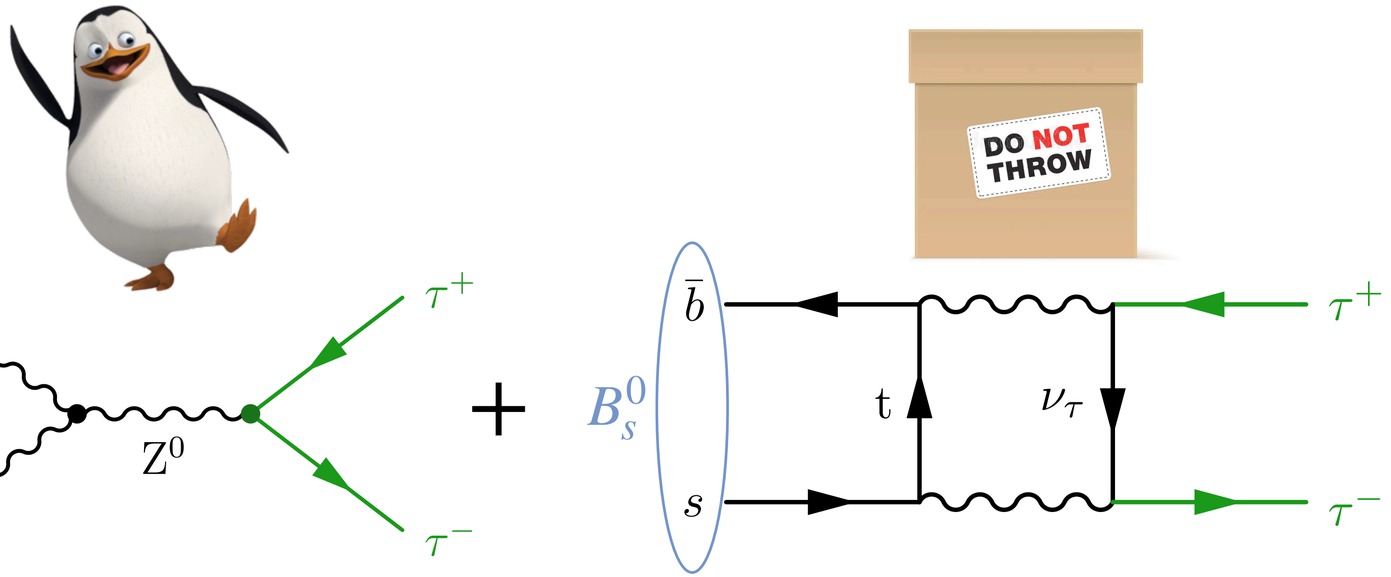


arXiv:2502.21223v4

IDEA detector concept

Physics of $B_s^0 \to \tau^+\tau^-$

- Hadronization fraction of $b-$quark into B_s^0 meson is $f_s = 11\,\%$
- $B_s^0 \to \tau^+\tau^-$ is a Flavour Changing Neutral Current **(FCNC) process**, dominated by two main Feynman diagrams: the penguin and box diagrams.
- It is a loop-induced process, which explain the low expected branching ratio



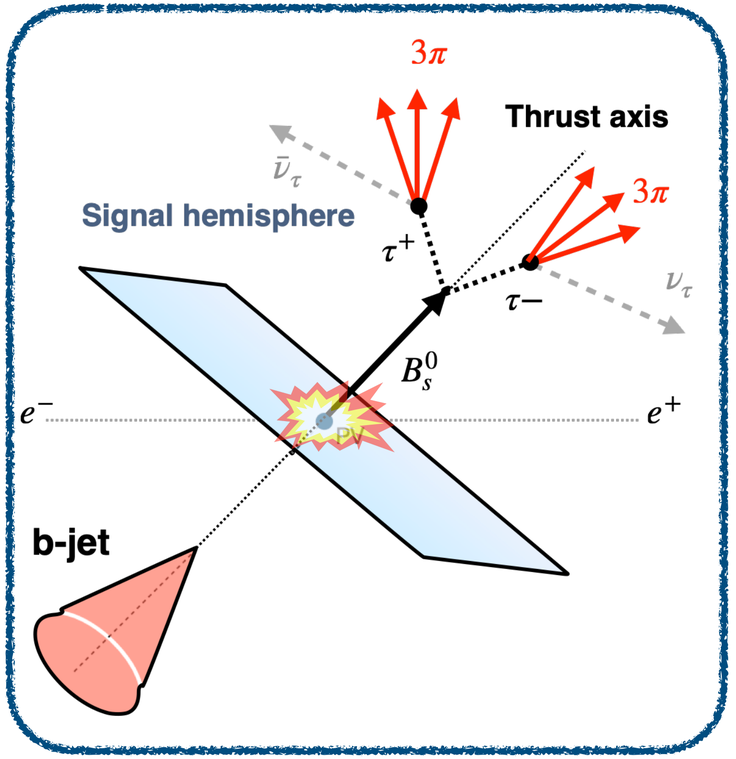
SM process

Physics of $B_s^0 \to \tau^+\tau^-$

Event Topology

- Collision at **center of mass energy 91 GeV** (Z^0pole)
- **Thrust axis reconstruction** is performed to characterize the event topology

$$T = \max \left[\frac{\sum_i |p_i \cdot n|}{\sum_i |p_i|} \right]$$



- The Thrust axis divides the event into two regions, referred as hemispheres

Monte Carlo samples

Signal, 10M events

- $B_s^0 \to \tau^+\tau^-, \tau \to 3\pi + \nu_\tau$

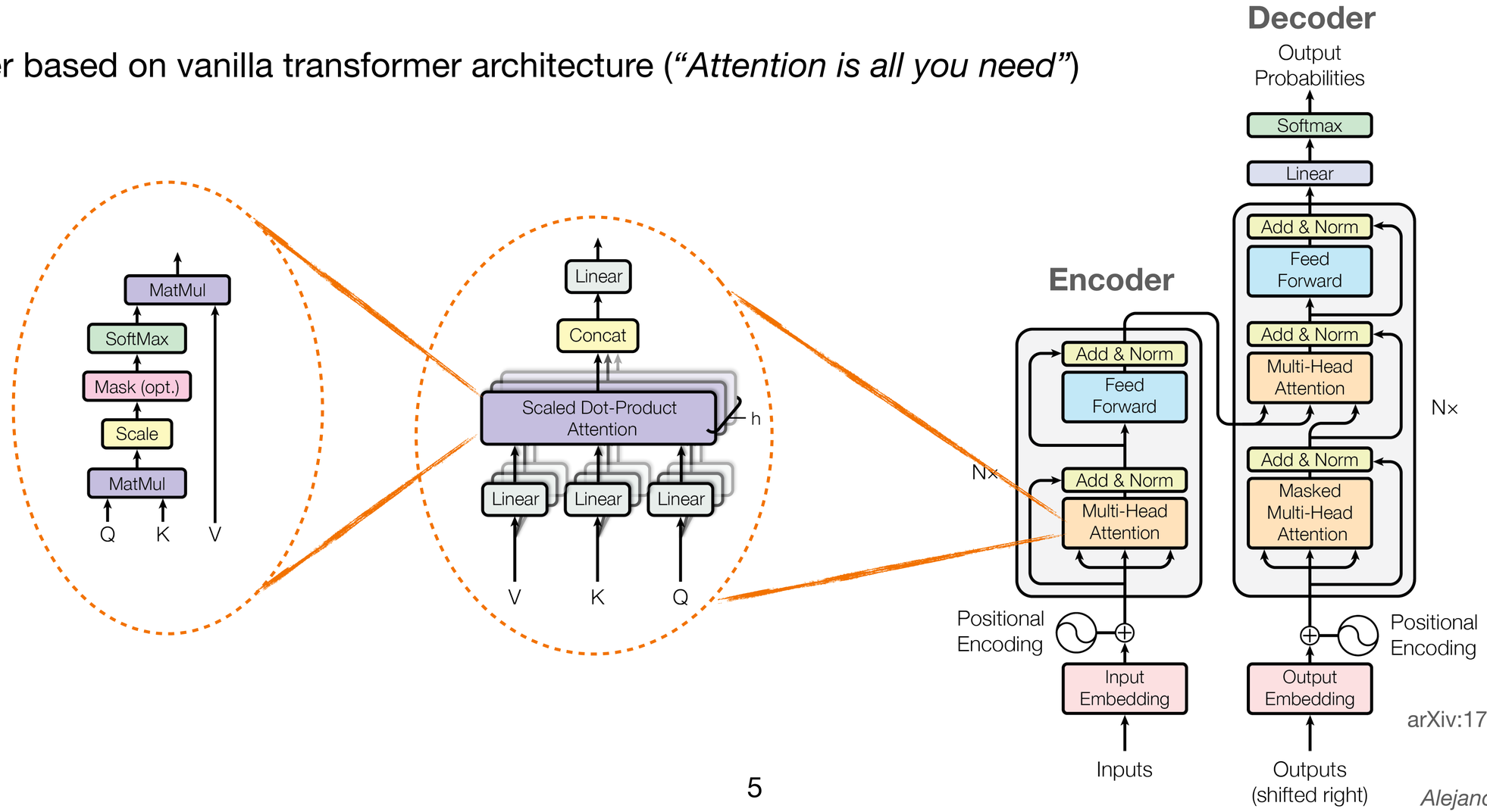
Background, 500M events for each process

- $Z^0 \to b\bar{b}$ inclusive $Z^0 \to s\bar{s}$ inclusive
- $Z^0 \to ud$ inclusive. $Z^0 \to c\bar{c}$ inclusive

DELPHES
fast simulation

Transformer - Based Classifier

- Secret of transformer is the self attention mechanism (Scaled Dot Product Attention)
- Classifier based on vanilla transformer architecture (*"Attention is all you need"*)



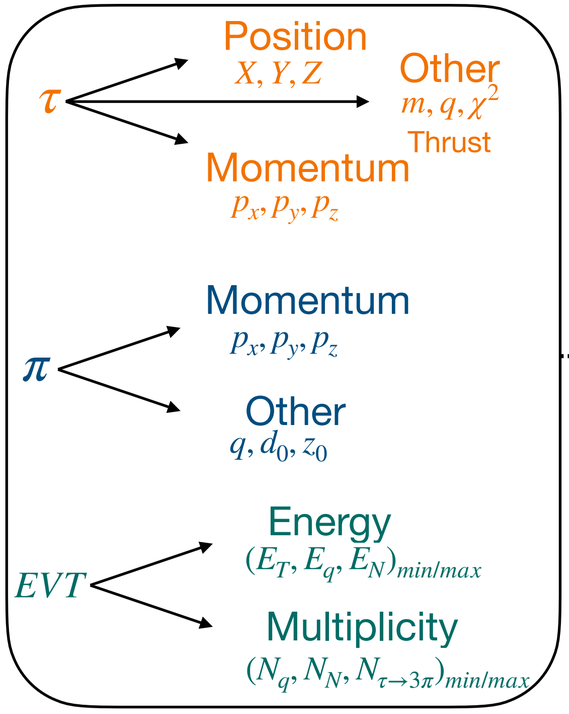
arXiv:1706.03762v7

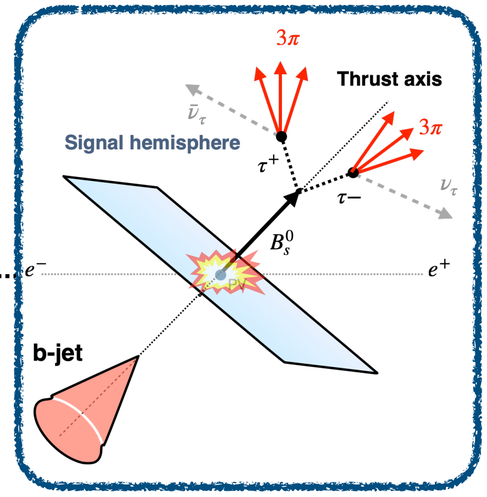
Transformer - Based Classifier

- **Goal:** Build a classifier to distinguish $B_s^0 \to \tau^+\tau^- \to 6\pi + 2\nu_\tau$ events vs $Z^0 \to b\bar{b}$ inclusive background

- **Input features:** Event level information and reconstructed τ and π features

Input variables list



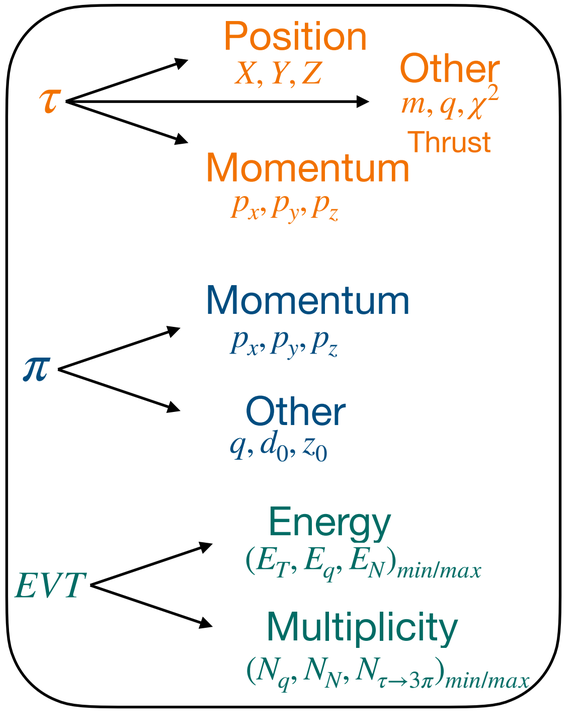


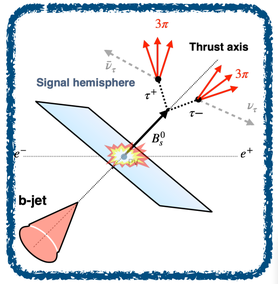
Transformer - Based Classifier

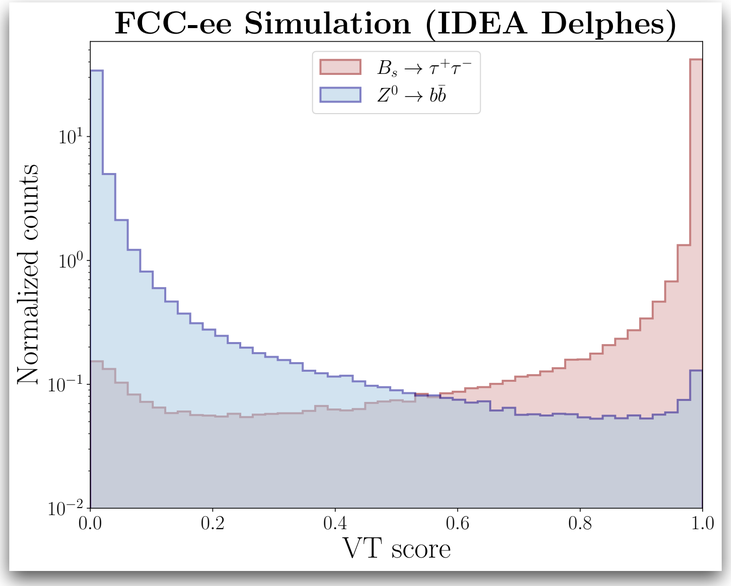
- **Goal:** Build a classifier to distinguish $B_s^0 \to \tau^+\tau^- \to 6\pi + 2\nu_\tau$ events vs $Z^0 \to b\bar{b}$ inclusive background

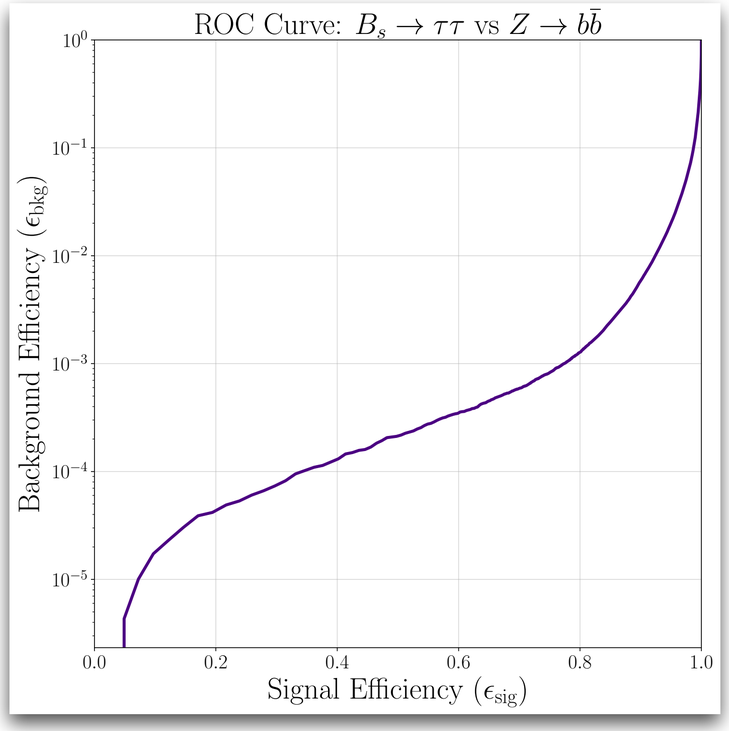
- **Input features:** Event level information and reconstructed τ and π features

Input variables list





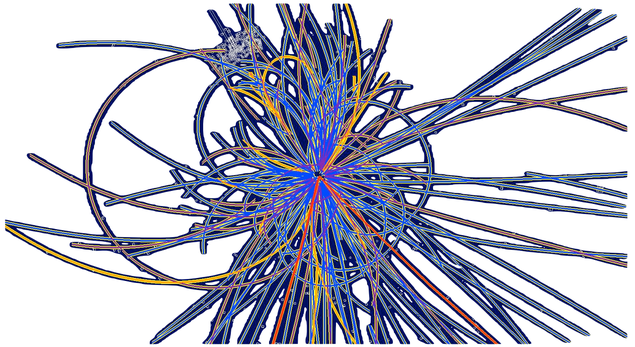




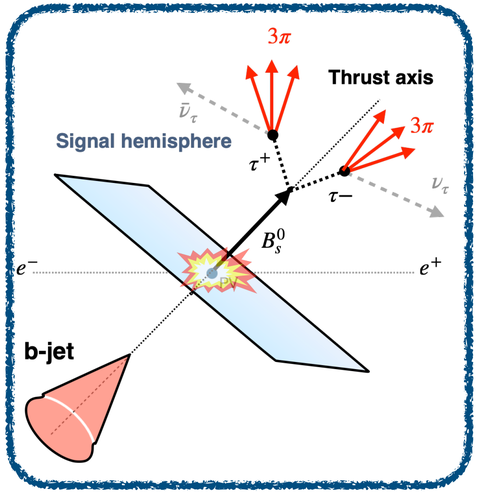
$\epsilon_{signal} \approx 80\,\%$ and $\epsilon_{bkg} \approx 0.1\,\%$

Transformer - Based Classifier

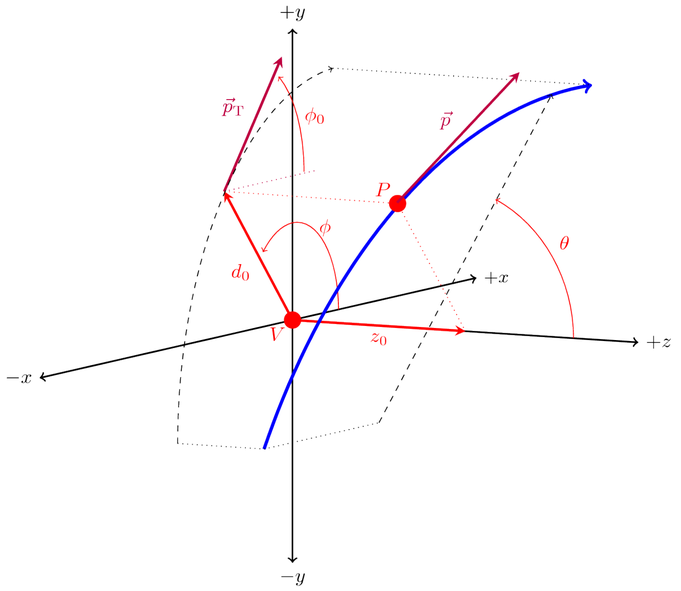
Track information



+



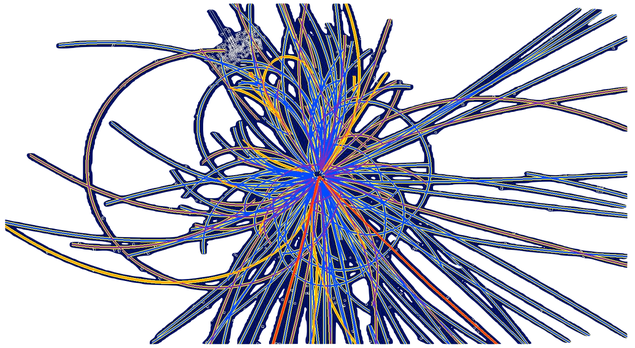
- Transformer architecture ideal to catch correlations with track information
- Helix parameters $(d_0, z_0, \phi_0, \omega, \tan(\lambda))$ of tracks (trajectory)



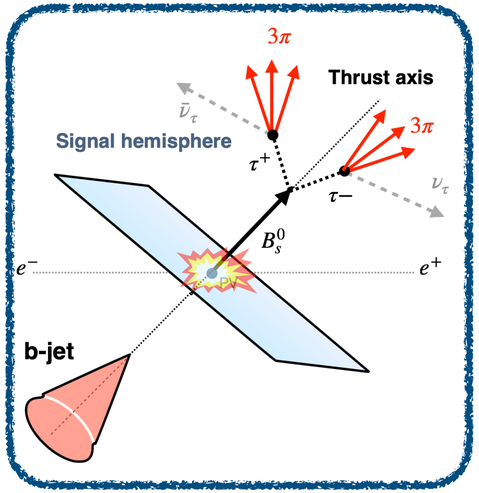
Transformer - Based Classifier

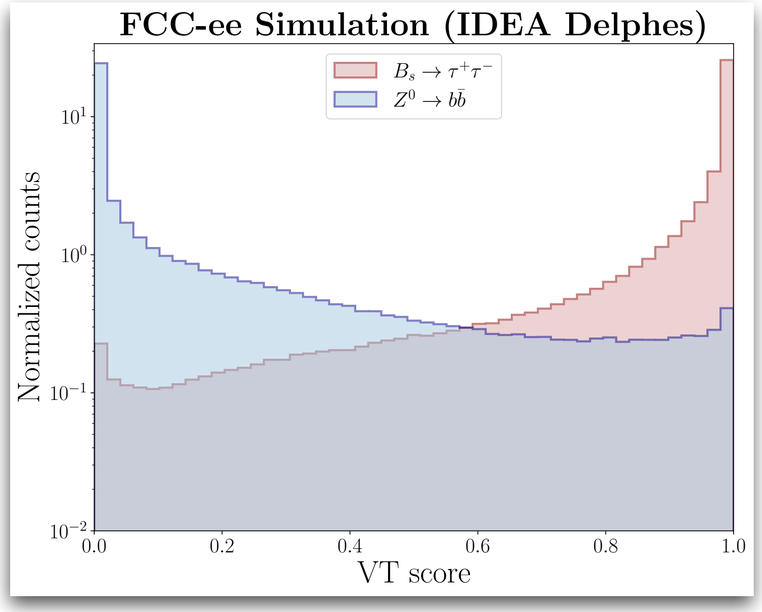
Track information

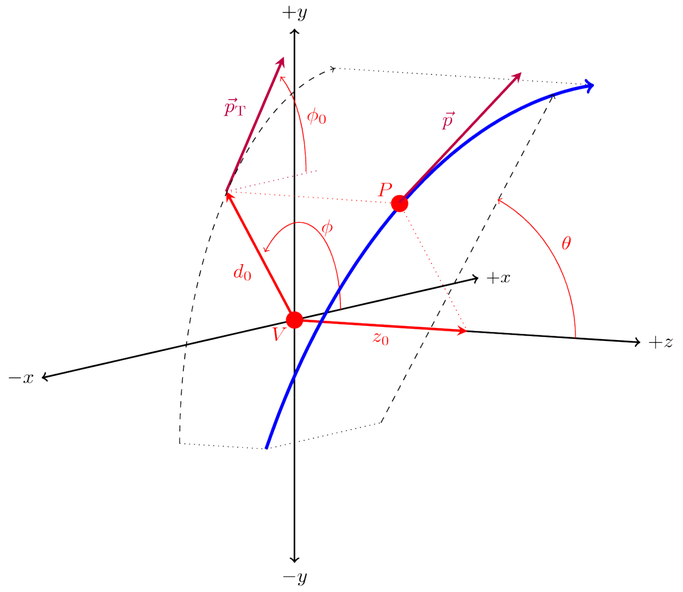
- Transformer architecture ideal to catch correlations with track information
- Helix parameters $(d_0, z_0, \phi_0, \omega, \tan(\lambda))$ of tracks (trajectory)

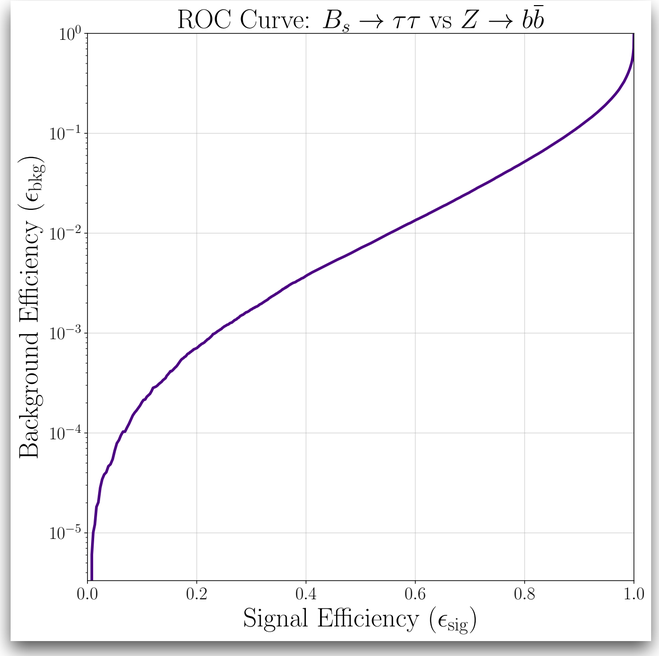


+









Conclusion and future plans

- Combination of tracks and event level information
- Optimizer selection and transformer implementation for $\tau \to 3\pi\nu_\tau$ channel.
- Signal and background yields extraction from efficiencies
- Extraction and report of upper limit for $B_s^0 \to \tau^+\tau^-$ with 95% or 68% CL within the FCC community
- Explore other of tau hadronic decay modes