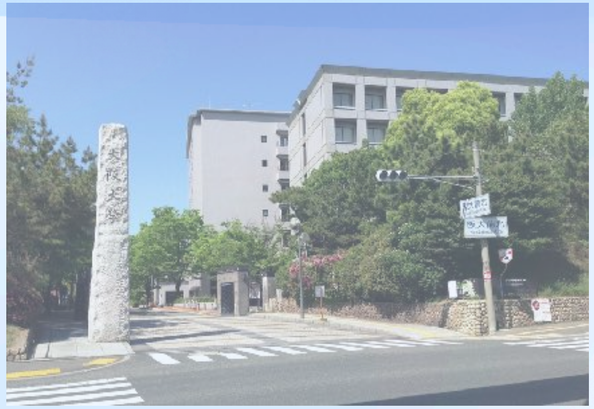
The 7th International Workshop on
"Higgs as a Probe of New Physics" (HPNP2025) @ University of Osaka
June 10, 2025

Deep learning for di- and tri-Higgs searches in multi-b channels



Cheng-Wei Chiang
National Taiwan University
National Center for Theoretical Sciences

CWC, Feng-Yang Hsieh, Shih-Chieh Hsu, and Ian Low, JHEP 09 (2024) 139 (2401.14198 [hep-ph]) — di-Higgs
CWC, Feng-Yang Hsieh, Shih-Chieh Hsu, Zhi-Zhong Li, and Ian Low, 2506.xxxxx [hep-ph] — tri-Higgs

Outline

- Motivations
- SPA-NET
- The 4-b channel
- The 6-b channel
- Summary

Motivations

Higgs Potential and Self Coupling

- To verify the EWSB mechanism, one needs to
  - measure the Higgs **couplings** to other SM particles; and
  - determine the Higgs **potential shape**, particularly through its **self coupling**.

no measurement yet!

- The Higgs potential (up to dim-4) can be generally parametrized as

$$V(h) = \frac{1}{2} m_h^2 h^2 + \lambda v h^3 + \lambda_{4h} h^4$$

  where the **trilinear** and **quartic** couplings in the SM are given by

$$\lambda^{\rm SM} = 4\lambda_{4h}^{\rm SM} = \frac{m_h^2}{2v^2}$$

- The determination of λ may tell us whether there could be a sufficiently strong **first-order EWPT** in the early Universe.

Kanemura, Okada, Senaha 2005

Di- and tri-Higgs Productions at LHC

- The **di-Higgs and tri-Higgs productions** at hadron colliders can in principle be used to probe the trilinear and quartic couplings.

for discussions of di-Higgs production, see also talks by Moretti, Braathen, and Heinemeyer

- They are, however, challenging measurements for the SM due to their **small** production cross section through GGF at the 13-TeV LHC:

$$\sigma(pp \to hh)^{\mathrm{GGF}}_{\mathrm{SM}} \sim 30 \text{ fb}$$
$$\sigma(pp \to hhh)^{\mathrm{GGF}}_{\mathrm{SM}} \sim 0.047 \text{ fb}$$

LHC Higgs Cross Section WG

- Current cross section limits are $\gtrsim 70$ fb on the di-Higgs search and $\gtrsim 60$ fb on the tri-Higgs search, depending on the channel.

ATLAS 2023, 2025

Multi-b Channels

- With a mass of 125 GeV, the Higgs boson predominantly decays into a $b\bar{b}$ pair.
  ➠ multi-b channels offer the largest rates, but suffer from large QCD background of multi-jet production
  ➠ partly solved by b-**tagging**, **top veto**, and **invariant mass cuts**

ATLAS 2019, 2022, 2023
CMS 2018, 2022

- Because of the inability to distinguish a b-jet from a $\bar{b}$-jet,* it is a challenging task to form the **correct pairing** among the b-jets to reconstruct the Higgs mass and the associated kinematic distributions.
  ➠ **combinatorially complex** for N essentially identical jets in traditional methods
  ➠ can we explore and preserve **permutation symmetries** in the process (e.g., multiple Higgs bosons and "indistinguishable" b-jets) to help simplify the calculations and thus **reduce the inference time**?

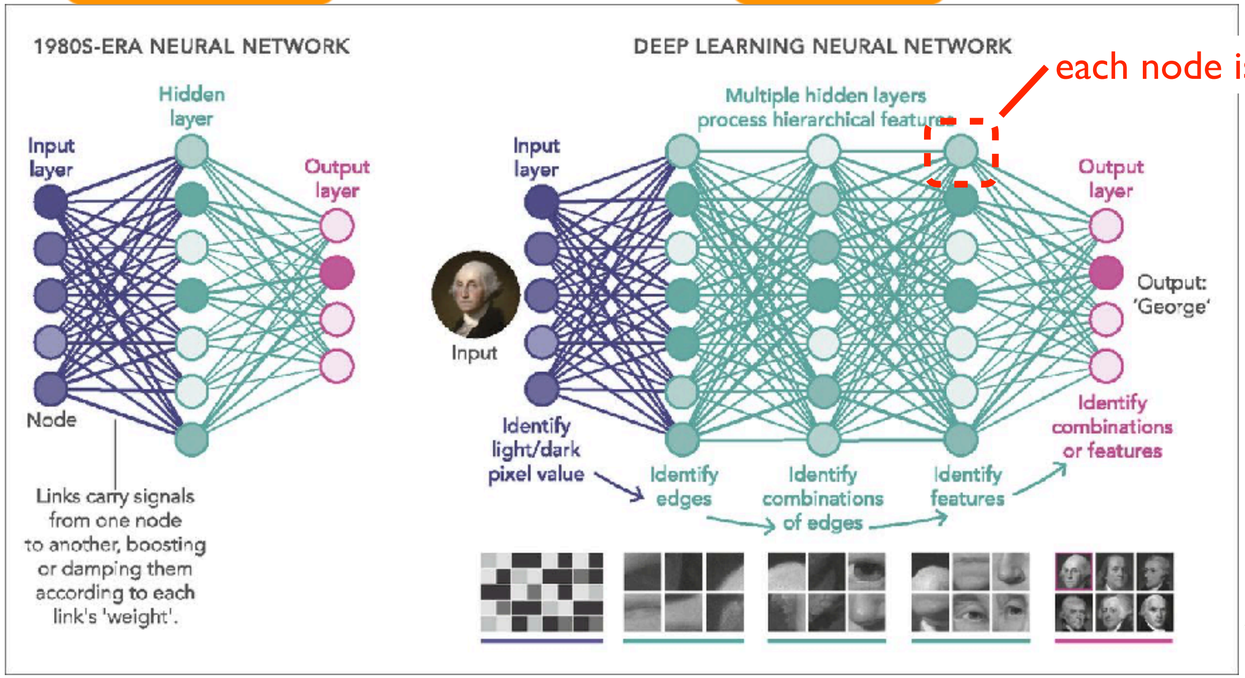
* Some experimental efforts are underway to tackle this issue.

Resurgence of Neural Networks

- Neural networks (NNs) have resurged in the last decade partly due to:
  - faster computers, with the **use of GPUs** versus the traditional use of CPUs,
  - better, **deeper algorithms** and NN architecture designs, and
  - increasingly **large datasets** being available for training.

shallow NN

deep NN



Waldrop 2019

Symmetry Preserving Attention Network (SPA-NET)

Symmetry Preserving Attention Network (SPA-NET)

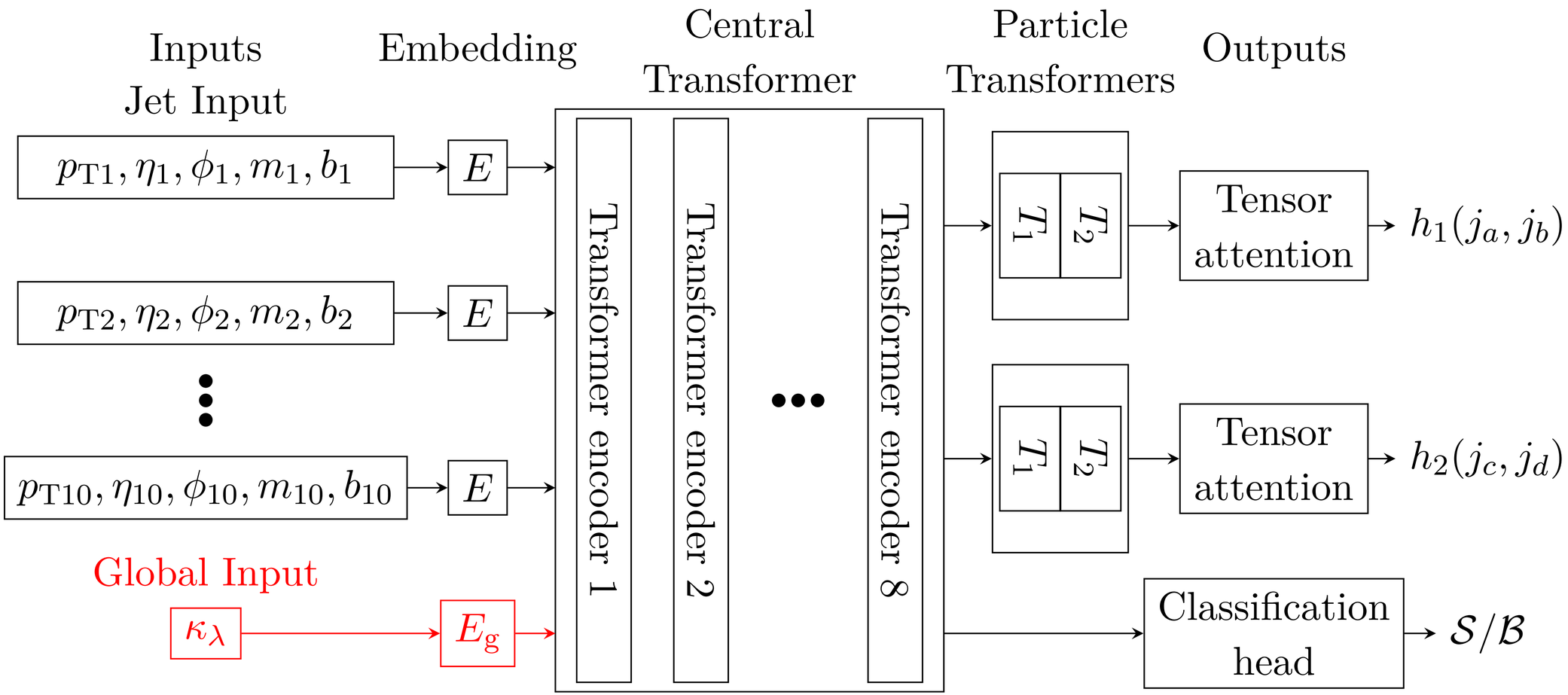
- SPA-NET is designed to solve such a problem: **reconstructing unstable particles (like top quarks or Higgs bosons) from the spray of jets** observed in detectors after proton-proton collisions.

Fenton et al. 2022
Shmakov et al. 2022
Fenton et al. 2024

- **Permutation-Free Assignment**: avoids explicit permutations by using a symmetry-aware loss function, but learns to match jets to particle decay products using an **attention mechanism**.
- **Attention-Based Design**: employs a **transformer-style attention mechanism** to model relationships between jets and particles.
- **Central and Particle Transformers**: includes a **central transformer** to capture **global event-level** information, and **particle-specific transformers** to refine the predictions for each candidate particle based on the global context.

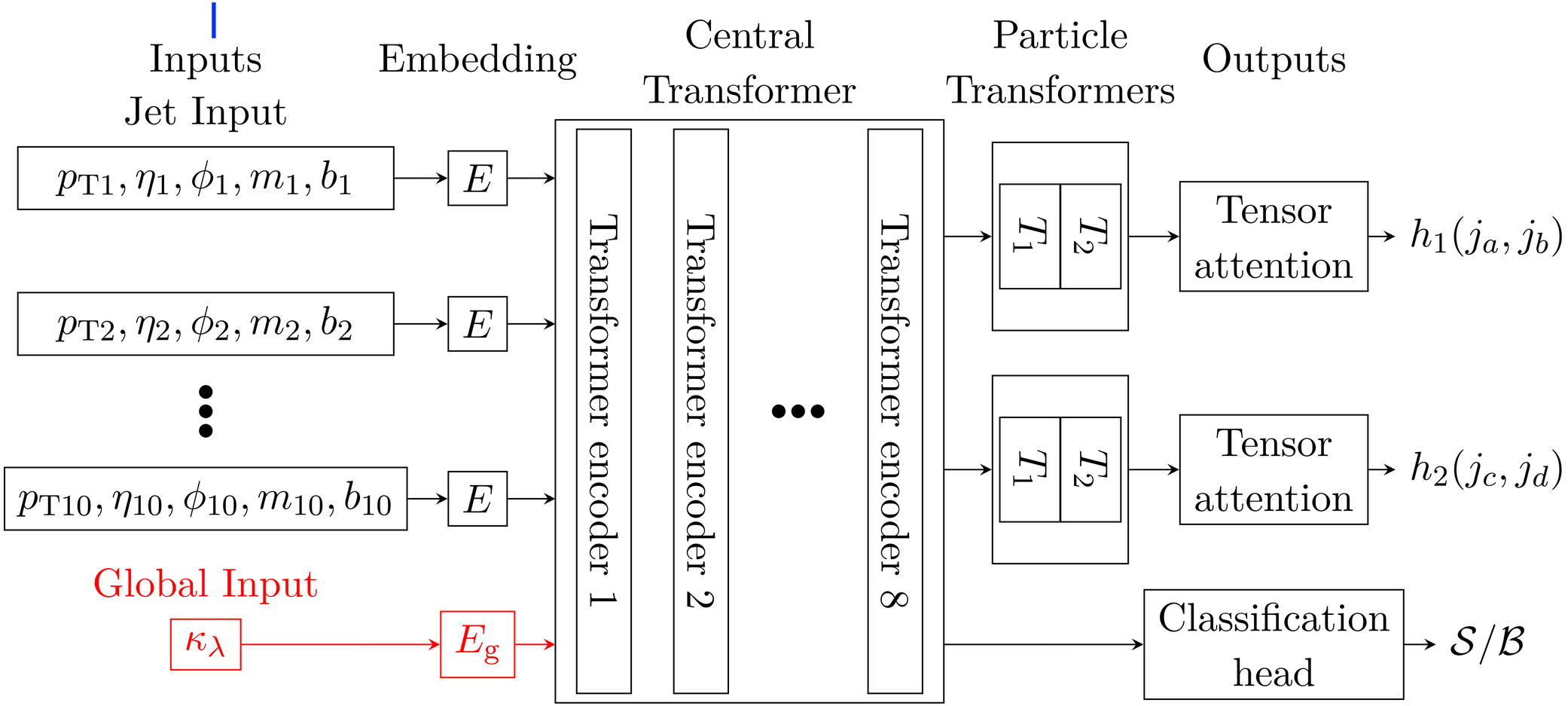
Model Structure of SPA-NET

- It achieves the tasks of **jet pairing/assignment** and **event classification**, taking into account **label** and **permutation symmetries** among the decay products.



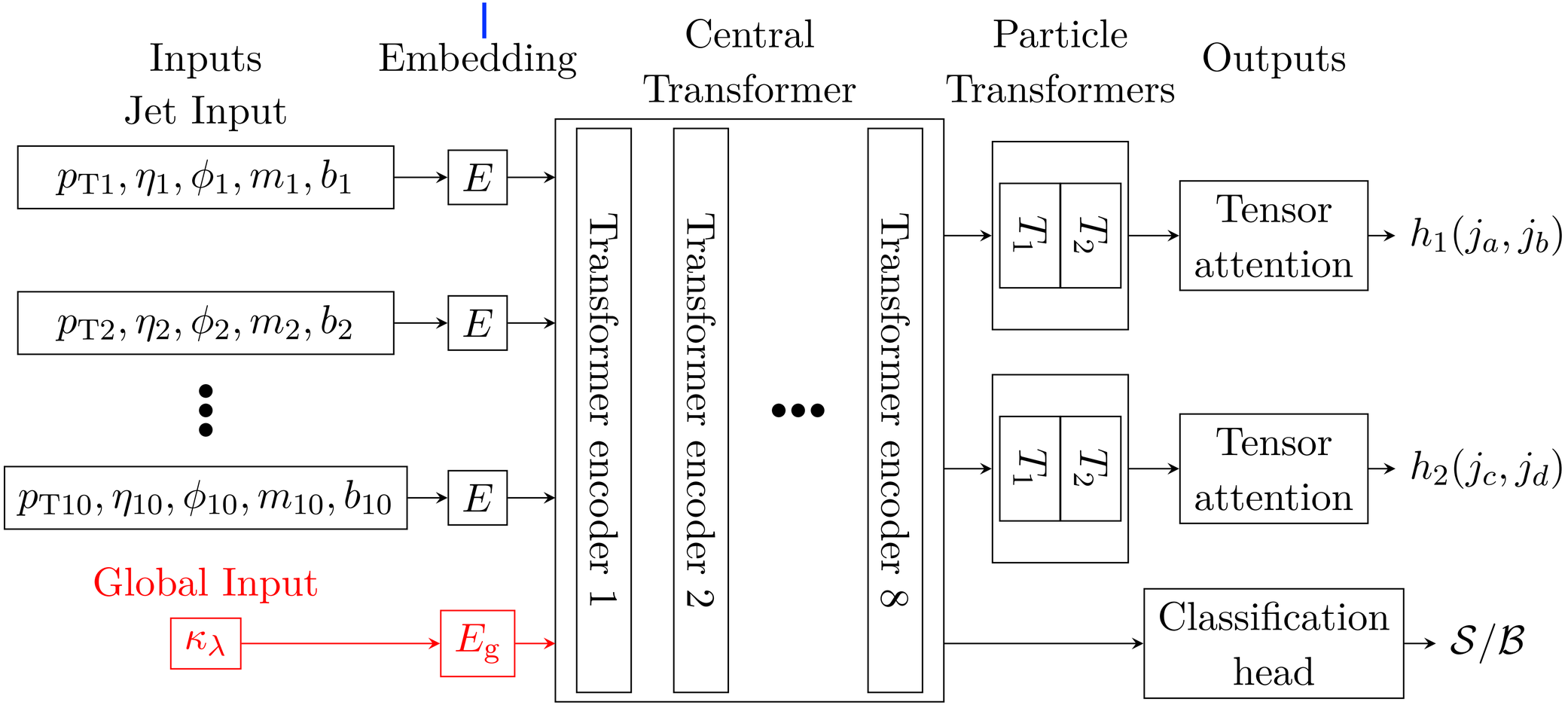
Model Structure of SPA-NET

A list of jets, each represented by its 4-component vector $(p_{\mathrm{T}}, \eta, \phi, m)$ as well as a boolean b-tag.



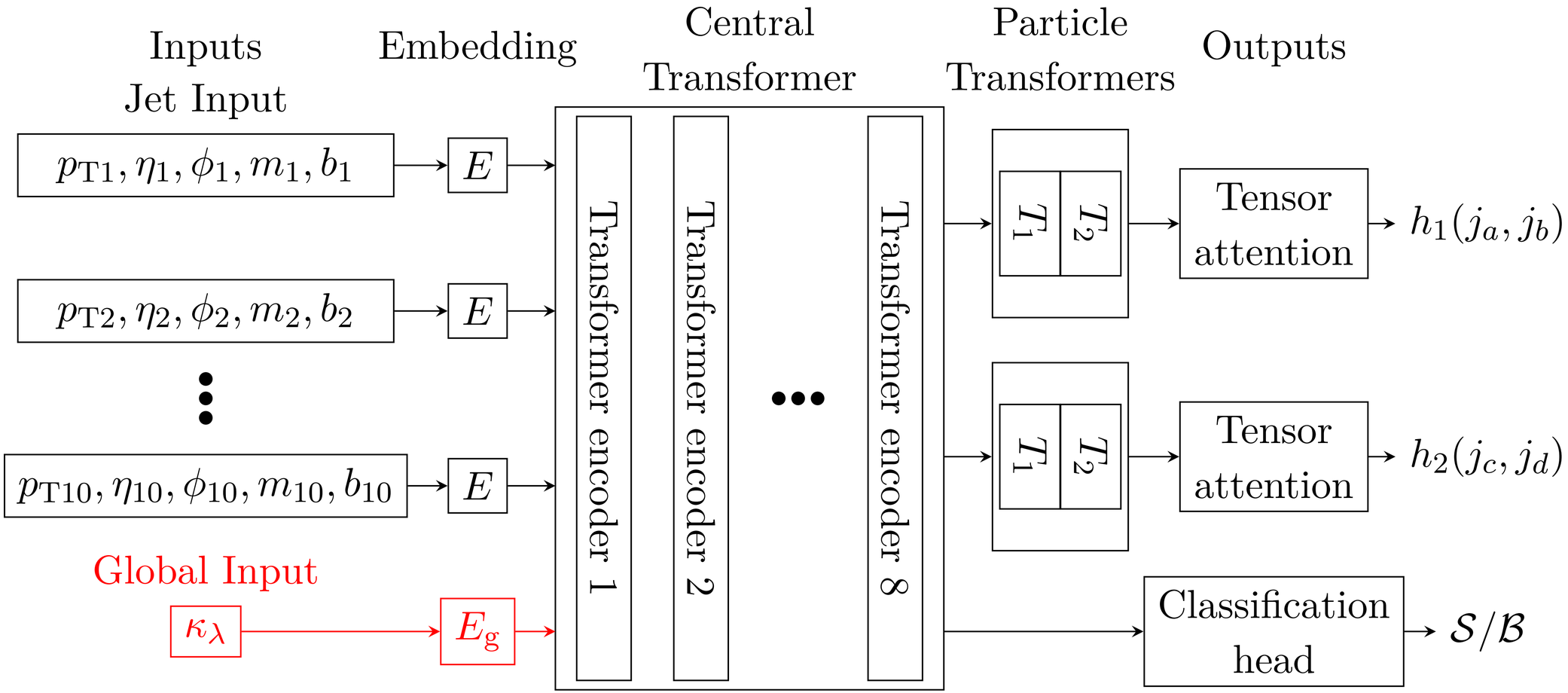
Model Structure of SPA-NET

The embedding blocks encode the input features to the embedding vectors living in the latent space.



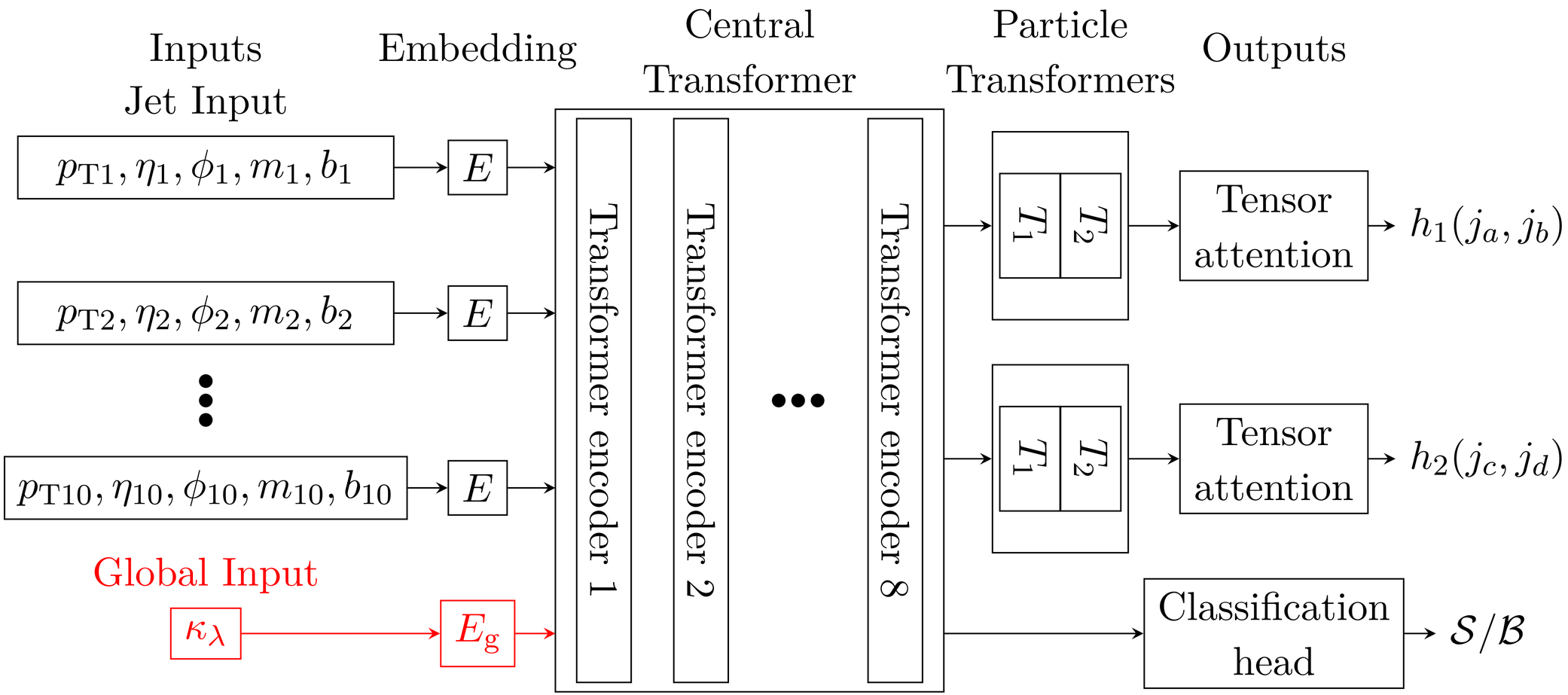
Model Structure of SPA-NET

With a stack of transformer encoders, it learns contextual information, encodes this information and outputs it into event embedding vectors for jet assignment and classification.



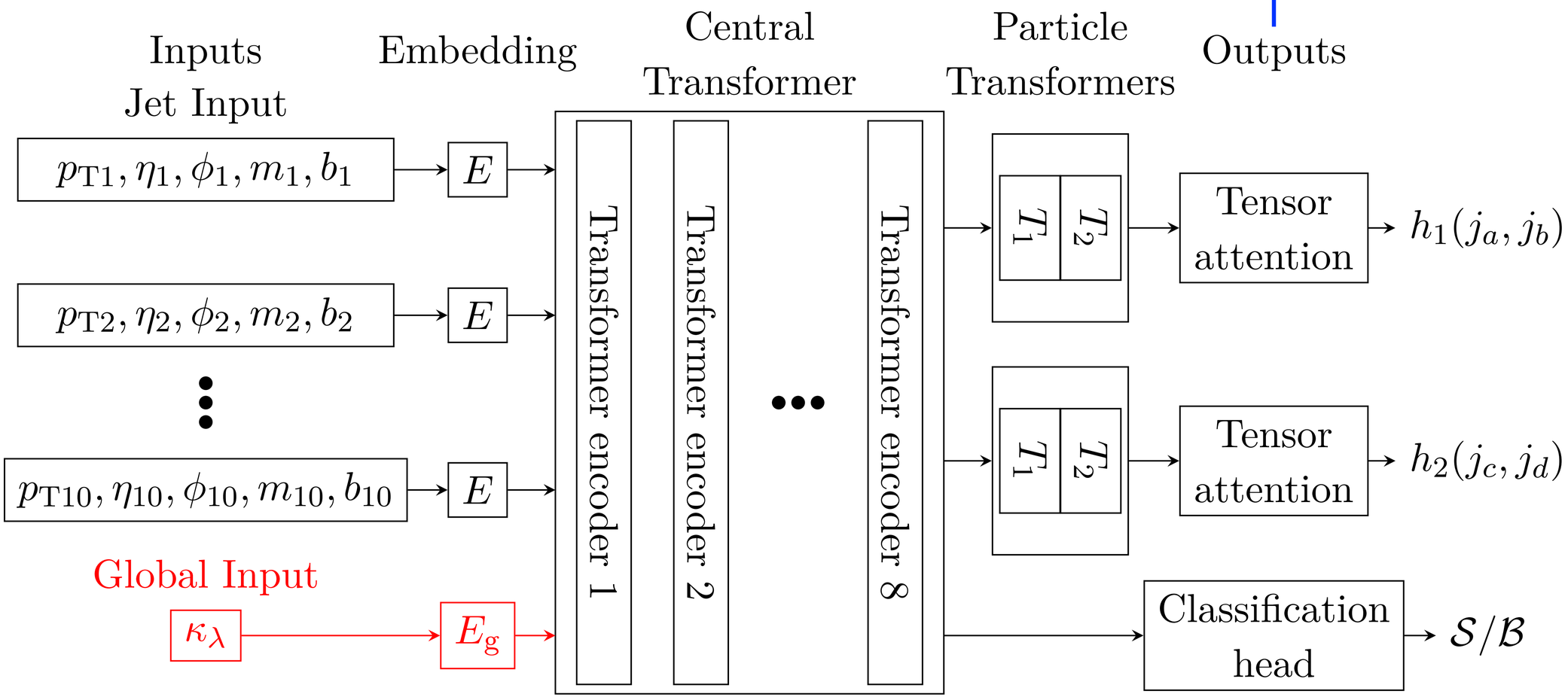
Model Structure of SPA-NET

It recognizes patterns in how particles are grouped in a jet and thus different types of jets, such as those produced by b's, W's, or h's, thus capture the particle-specific information



Model Structure of SPA-NET

It performs a weighted scalar product of the transformer-encoded jet vectors, outputting a tensor that gives essentially the probability. The technique of symmetric tensor attention is employed to construct a tensor with permutation symmetries of labels (e.g., the $b\bar{b}$ pairs).



Event Generation

- Use **`MadGraph5_aMC@NLO`** to generate both signal and background events at the LO partonic level for $\sqrt{s} = 13$ TeV with the **`NNPDF23_nlo_as_0119`** PDF.
- Use **`Pythia`** with ISR and FSR at the parton level and **`NNPDF2.3 LO PDF`** for parton showering and hadronization.
  ➠ there might be more than 4 or 6 jets after the jet clustering
- Use **`Delphes`** for fast detector simulation, with jet reconstruction by **`FastJet`** using the anti-k_t algorithm with a jet radius of $R = 0.4$, $p_{\mathrm{T}} > 40$ GeV and $|\eta| < 2.5$ and a basic 4 or 6 tag cut requiring at least 4 or 6 b-tagged jets.
- The decays $H \to hh, h \to b\bar{b}$ are implemented by **`MadSpin`**.

The 4-b Channel

Two Scenarios for di-Higgs production

- **On-shell**, **resonant** production of a BSM new scalar particle, which subsequently decays into a pair of Higgs bosons ($H \to hh$).
  ➠ as an explicit example, use the **2HDMs in the alignment limit**, where the Higgs pair is produced through the decay product of the heavy CP-even scalar, with no modification to the SM contribution

  Low, Shah, Wang 2022
  Chen, CWC, Low 2022

- **Non-resonant** production, where we rescale the SM trilinear coupling $\lambda = \kappa_\lambda \lambda^{\mathrm{SM}}$ (presumably due to corrections by new physics at a higher mass scale), of which the production cross section is a function, and assume no direct production of new particles.

Pairing Methods

- To reconstruct the two Higgs boson candidates, we consider **three** different methods:
  - The ΔR+**min-**D_{hh} **pairing**;

ATLAS 2019
ATLAS 2023

$$\left.\begin{aligned} \frac{360\ \text{GeV}}{m_{4\text{j}}} - 0.5 < \Delta R_{\text{jj}}^{\text{lead}} < \frac{653\ \text{GeV}}{m_{4\text{j}}} + 0.475 \\ \frac{235\ \text{GeV}}{m_{4\text{j}}} < \Delta R_{\text{jj}}^{\text{subl}} < \frac{875\ \text{GeV}}{m_{4\text{j}}} + 0.35 \end{aligned}\right\} \text{if } m_{4\text{j}} < 1250\ \text{GeV}$$

$$\left.\begin{aligned} 0 < \Delta R_{\text{jj}}^{\text{lead}} < 1 \\ 0 < \Delta R_{\text{jj}}^{\text{subl}} < 1 \end{aligned}\right\} \text{if } m_{4\text{j}} > 1250\ \text{GeV}$$

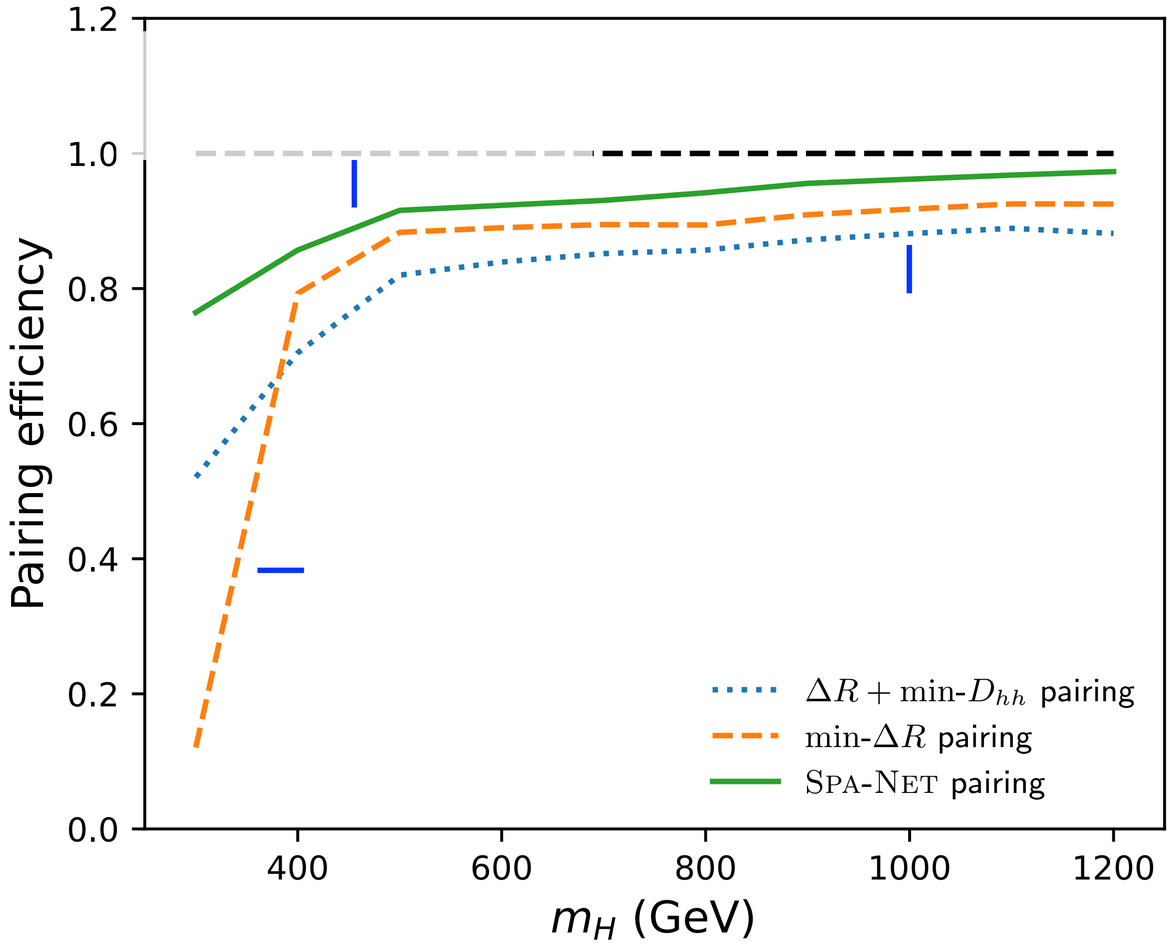
$$D_{hh} = \frac{\left| m_{h_1} - \frac{120}{110} m_{h_2} \right|}{\sqrt{1 + \left(\frac{120}{110}\right)^2}}$$

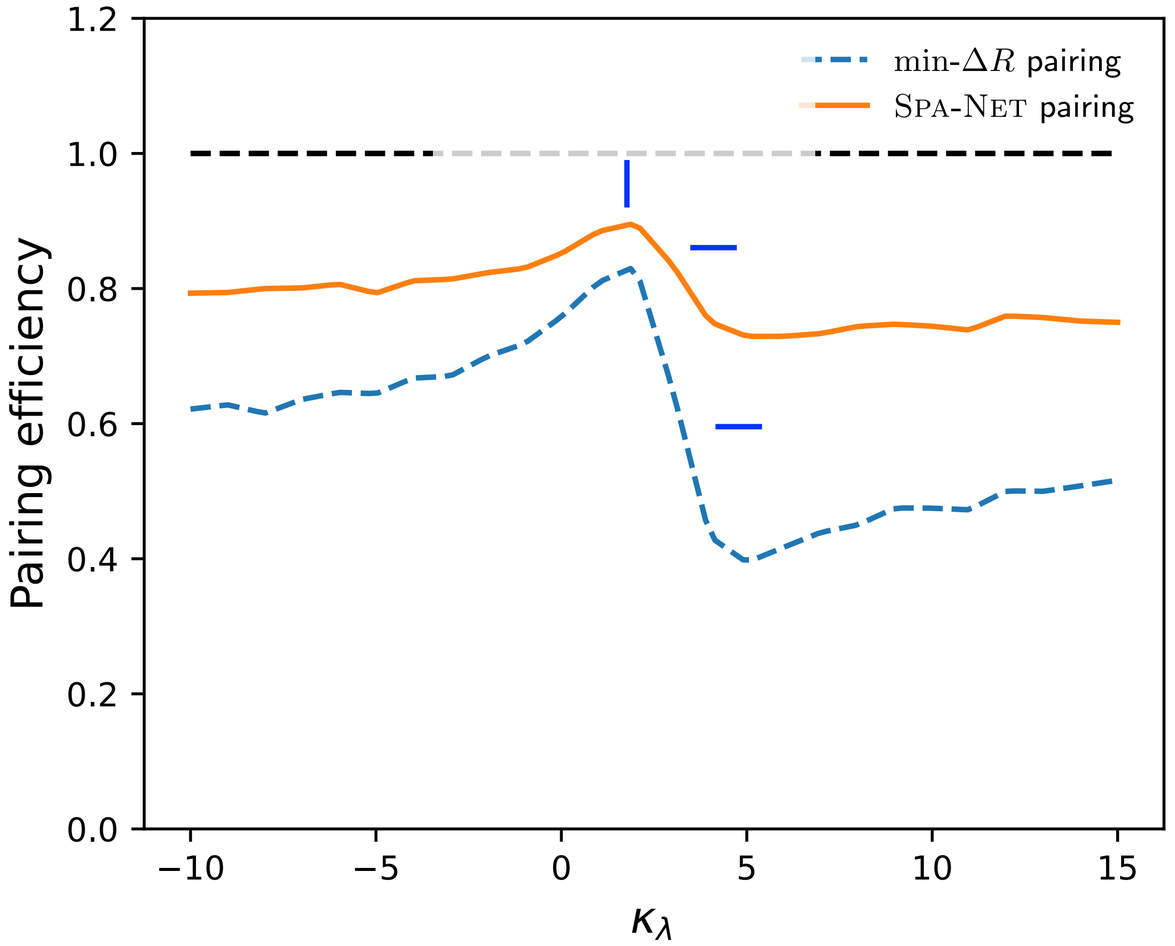
both are cut-based and inspired by the ATLAS analyses

  - The **min-**ΔR **pairing**;
    - Choose 4 b-tagged jets with the highest p_{T} , and select the pairing where the higher-p_{T} jet pair has the smallest ΔR separation.
  - The **SPA-NET**.

Jet Pairing Comparison

- SPA-NET pairing **outperforms** other methods for all mass and κ_λ values.





Neural Network (NN) Classifiers

- After jet pairing, one can also discuss **event classification** (signal vs background).
- In addition to the cut-based approach, two types of NN classifiers are considered in this study:
  - **Dense-NN**, a conventional deep machine learning architecture used as the **baseline** NN approach, and
  - **SPA-NET**, a novel deep NN that can do jet pairing and event classification.
- While the Dense-NN is trained **only** on the classification task, SPA-NET is trained on both **jet pairing** and **event classification**.
  ➠ training on both tasks allows SPA-NET to achieve better performance (owing to the help of jet assignment with building a **better embedding space structure**) compared to only training on the classification head.

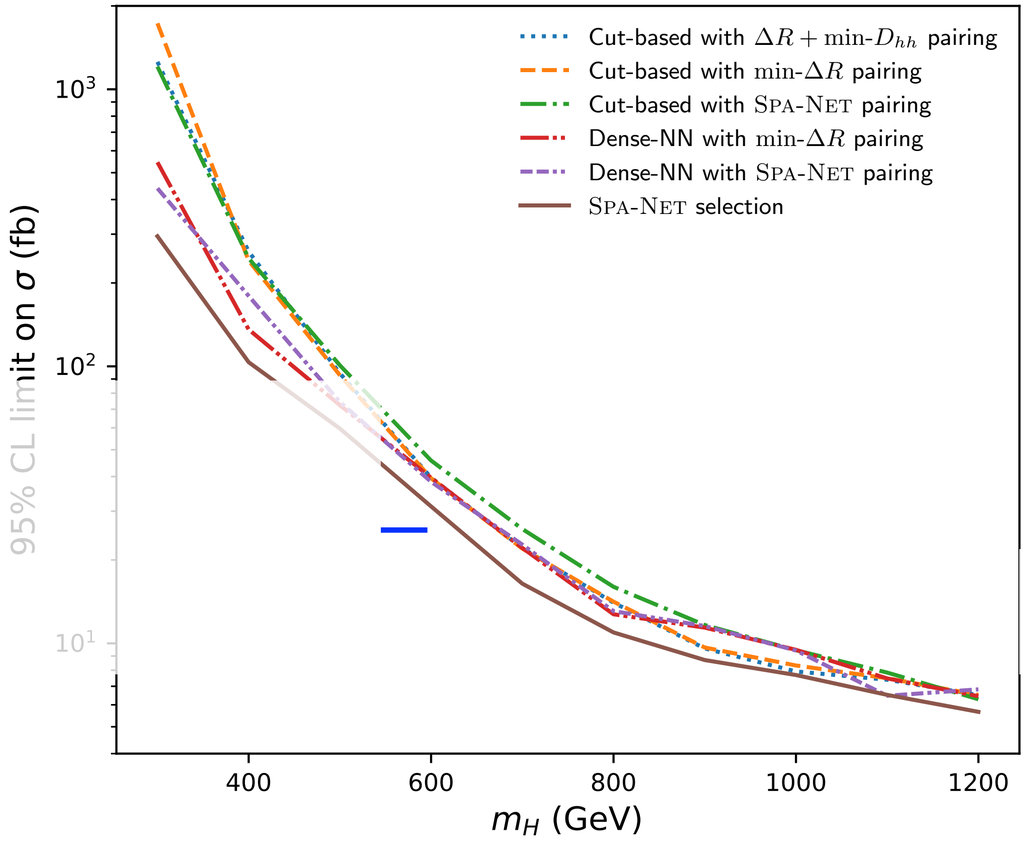
Di-Higgs Event Classification Comparison

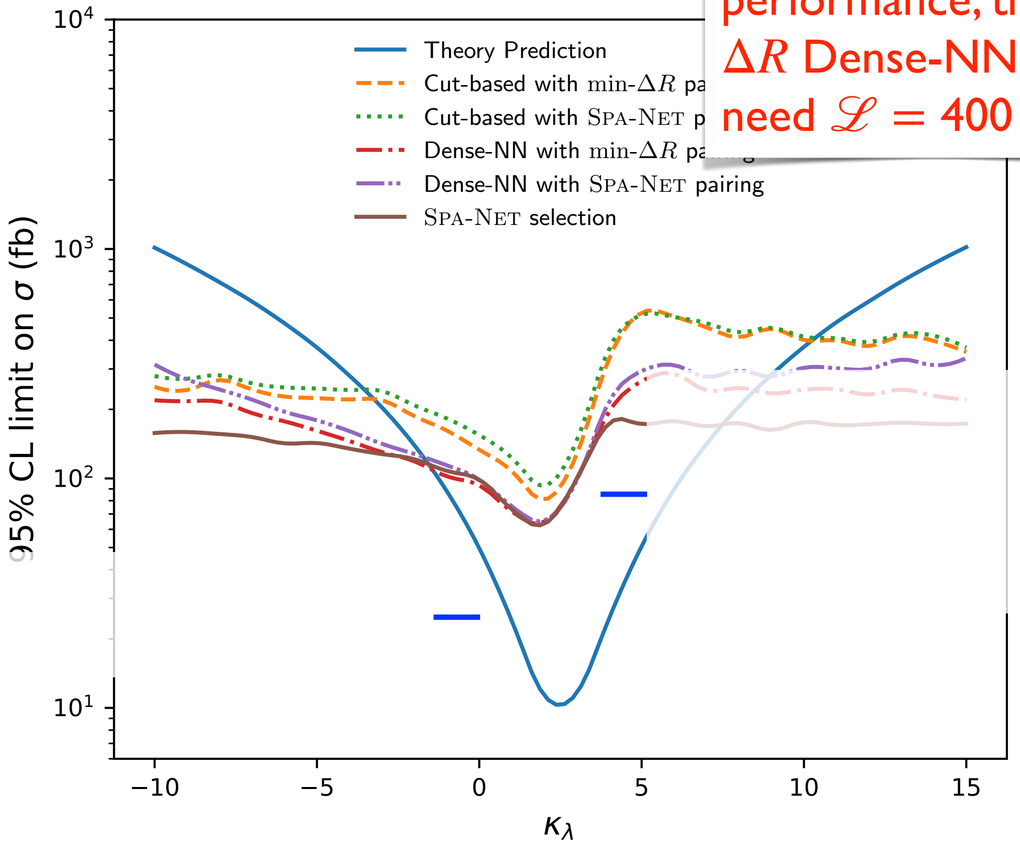
- SPA-NET has the **best performance** among the three event classifiers.

Classifier	ACC	AUC
Dense-NN with min-ΔR pairing	0.799 ± 0.011	0.881 ± 0.012
Dense-NN with SPA-NET pairing	0.803 ± 0.004	0.884 ± 0.004
SPA-NET	0.828 ± 0.002	0.911 ± 0.001

Upper Limits on Cross Sections

- Assume $\mathscr{L} = 300\ \text{fb}^{-1}$ for the 14-TeV LHC.
- SPA-NET has the **most superior performance** and provide the **most stringent constraints** throughout the considered m_H and κ_λ ranges.

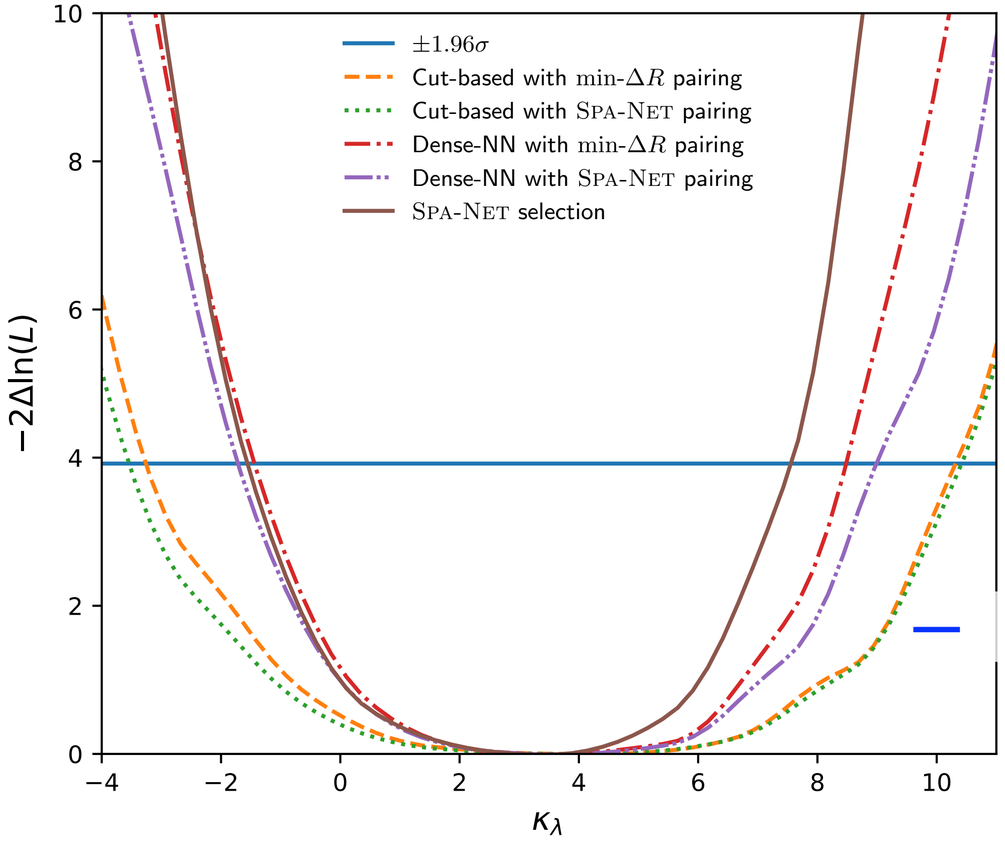




To achieve similar performance, the min-ΔR Dense-NN, would need $\mathscr{L} = 400\ \text{fb}^{-1}$.

Constraints on κ_λ

- Assume $\mathscr{L} = 300\ \text{fb}^{-1}$ for the 14-TeV LHC.
- SPA-NET provides the strongest constraints on κ_λ: beyond $[-1.56, 7.57]$ would be excluded at the 95% CL (1.96σ) level.



The 6-b Channel

Tri-Higgs Search

- We further demonstrate the power of SPA-NET in the tri-Higgs search at the LHC.
- The signal process is the **resonant** production of **three** $h_1(125)$ via GGF in the **two real singlet model (TRSM)**: Papaefstathiou, Robens, and Tetlalmatzi-Xolocotzi 2021

$$gg \to h_3 \to h_2 h_1 \to h_1 h_1 h_1$$

where we consider **five benchmark mass points**: ATLAS 2025

(m_{h_3}, m_{h_2}) [GeV]	$(420, 280)$	$(500, 275)$	$(500, 300)$	$(520, 325)$	$(500, 350)$
σ [fb]	80.20	63.98	60.79	50.43	46.15

where the cross sections are computed for the 13-TeV LHC with generation-level cuts: $|\eta_b| < 2.6$ and $p_T^b > 24$ GeV.

- The dominant irreducible background is QCD $pp \to b\bar{b}b\bar{b}b\bar{b}$. Others involving light-flavor or charm jets being misidentified as b-jets are found to contribute less than 10% of the total background and are neglected.

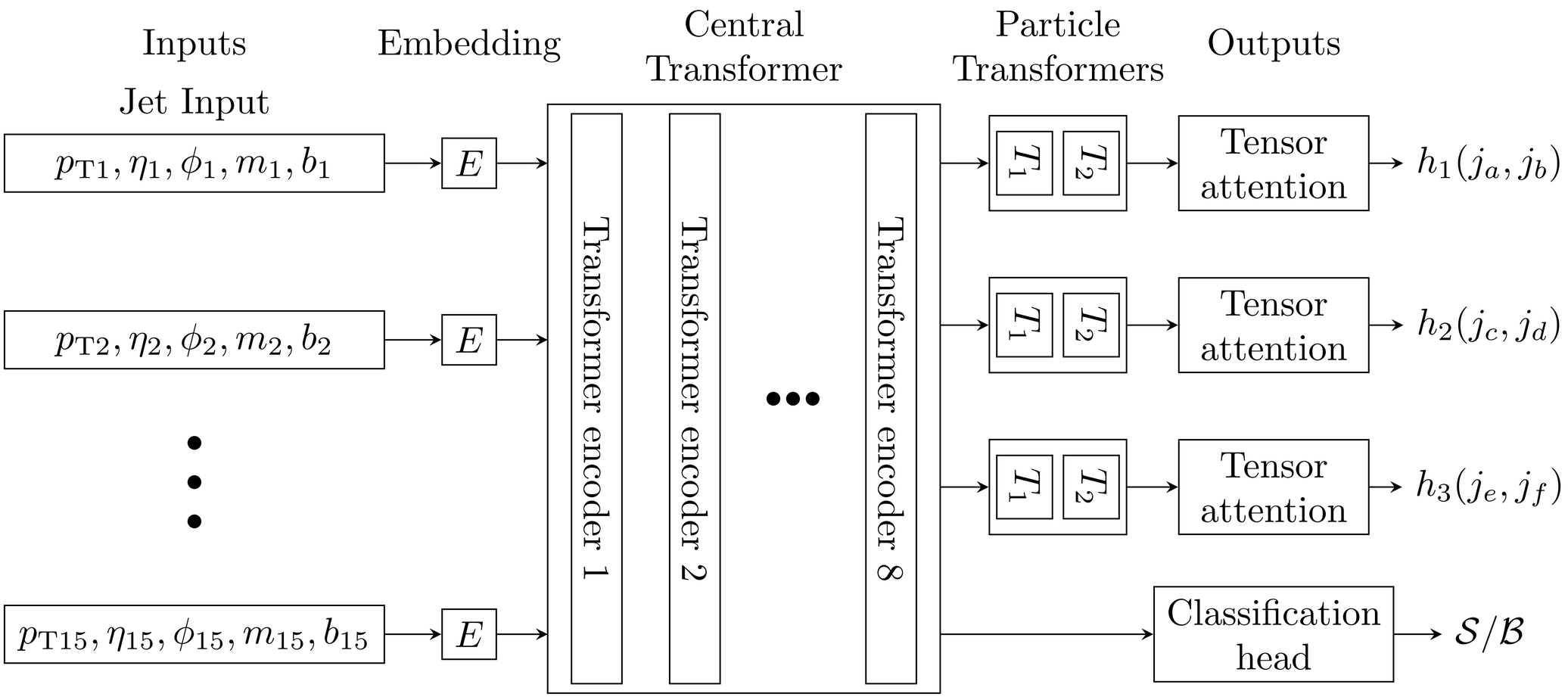
Jet Pairing

- Three jet pairing strategies considered and compared in this work:
  - χ^2 **pairing**: Selects the configuration of three Higgs boson candidates that minimizes $\chi^2 = \left(m_{h1} - m_h\right)^2 + \left(m_{h2} - m_h\right)^2 + \left(m_{h3} - m_h\right)^2$, where m_{hi} is the invariant mass of the i-th Higgs candidate and $m_h = 125$ GeV.

ATLAS 2025

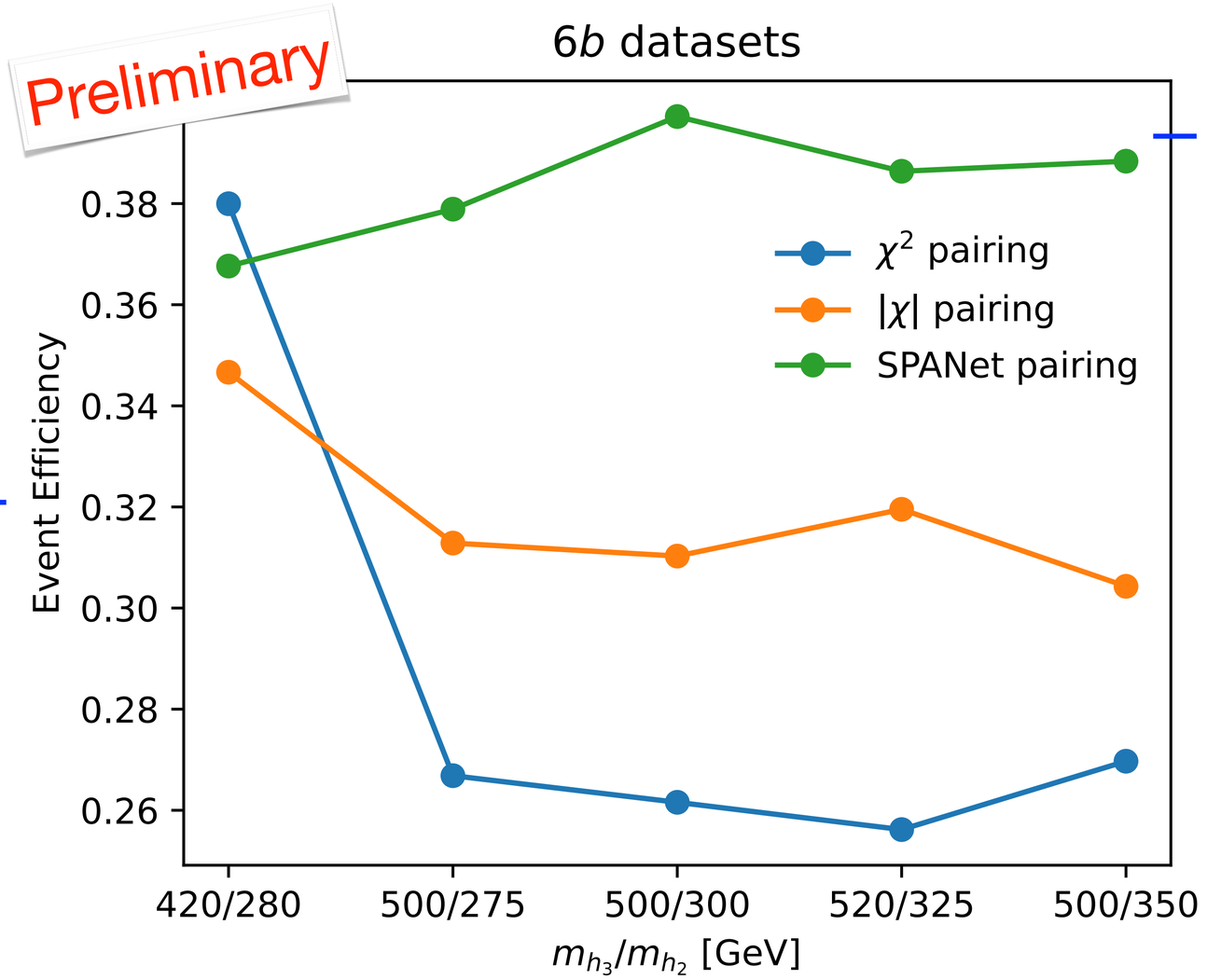
  - $|\chi|$ **pairing**: Selects the configuration of three Higgs boson candidates that minimizes $|\chi| = |m_{h1} - 120| + |m_{h2} - 115| + |m_{h3} - 110|$, where $(120,115,110)$ GeV are the peaks of the leading, sub-leading, and least-leading candidate Higgs invariant mass distribution, respectively.
  - **SPA-NET pairing**: Use a deep learning NN that predicts optimal jet assignments using **jet kinematics** and **symmetry-preserving attention mechanisms**.

Model Structure of SPA-NET

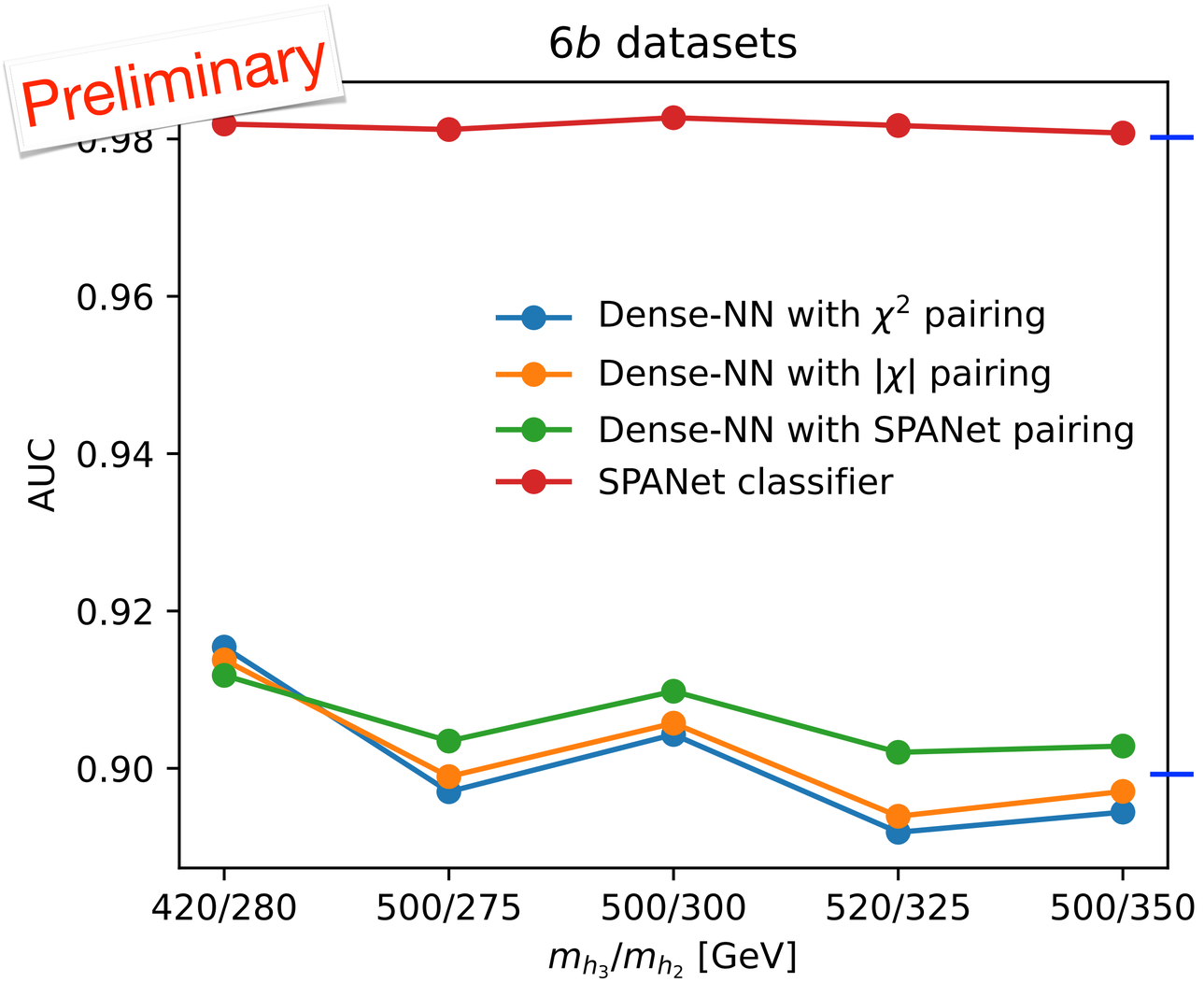


Jet Pairing Comparison

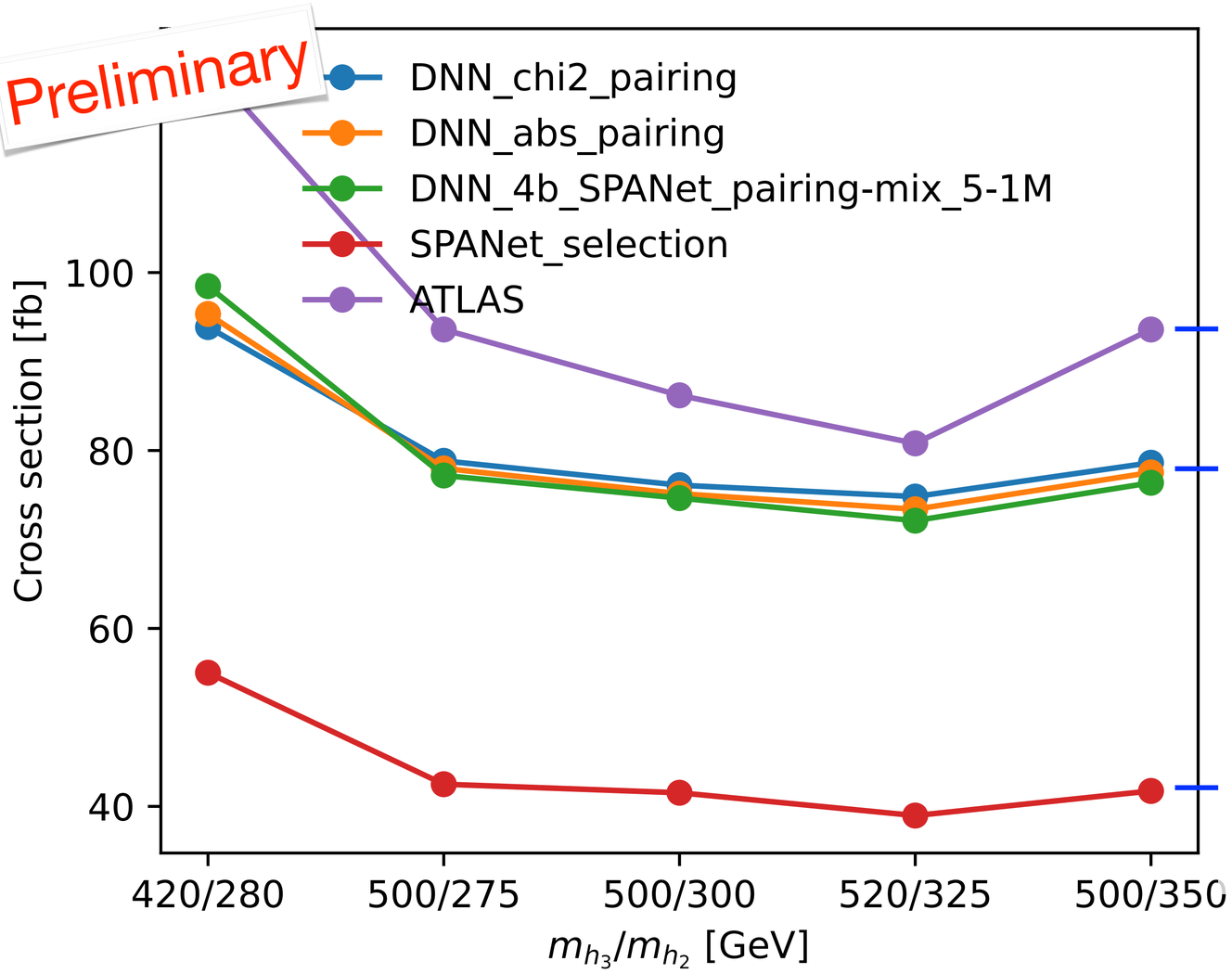
fraction of events where all three Higgs bosons are correctly reconstructed via jet pairing



Tri-Higgs Event Classification Comparison



Upper Limits on Cross Sections



ATLAS 2025

Summary

- We propose a novel deep neural network architecture called **SPA-NET** to improve the analysis for di- and tri-Higgs events in the $4b$ and $6b$ final states, respectively.
- We show that SPA-NET outperforms traditional cut-based and χ^2 methods in **pairing efficiency** and the Dense-NN classifier in **classification**.
  ➠ more stringent limits on the **production cross sections** and the **Higgs self coupling**
- Our results highlight the effectiveness of the task-specific design of SPA-NET in addressing the **combinatorial challenges** of multi-jet final states and motivate its applications to other complex hadronic final states at the LHC.

Thank You!