Improving Multi-Higgs sensitivity in the hadronic final state using ML

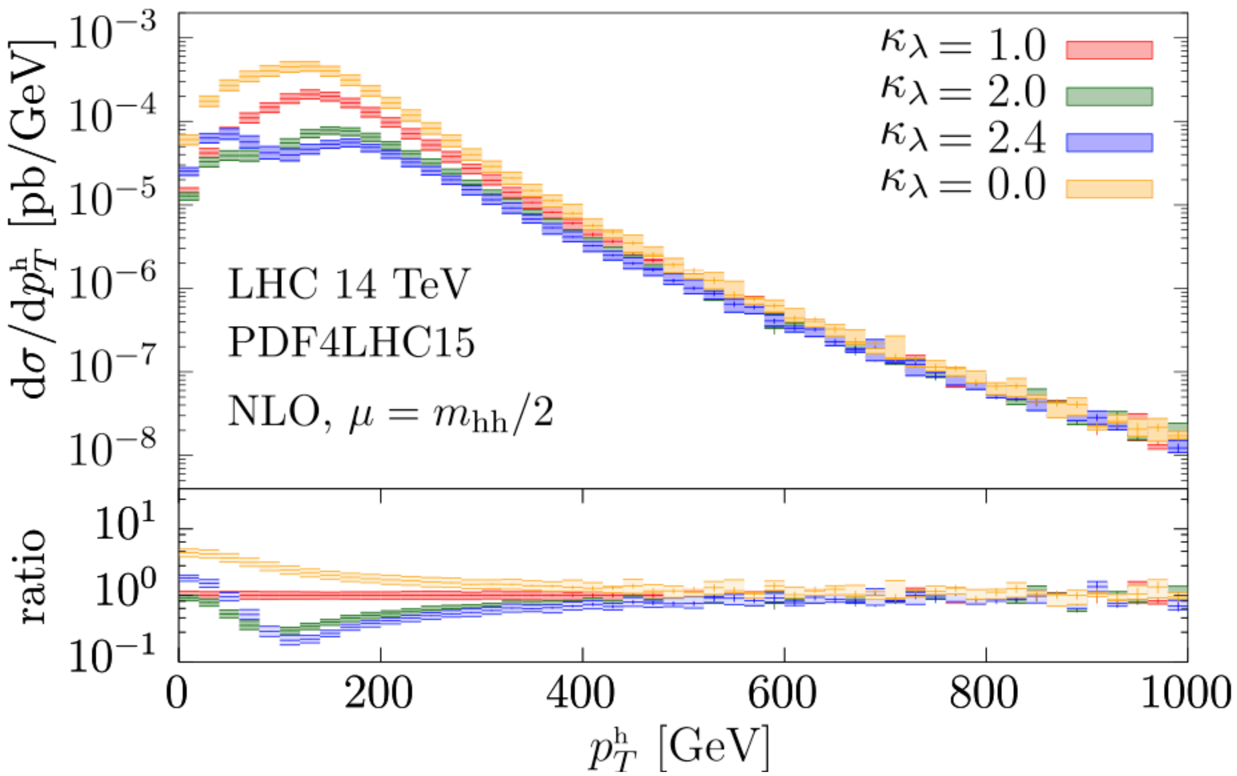
Daniel Diaz, Javier Duarte, Sanmay Ganguly, Brian Sheldon

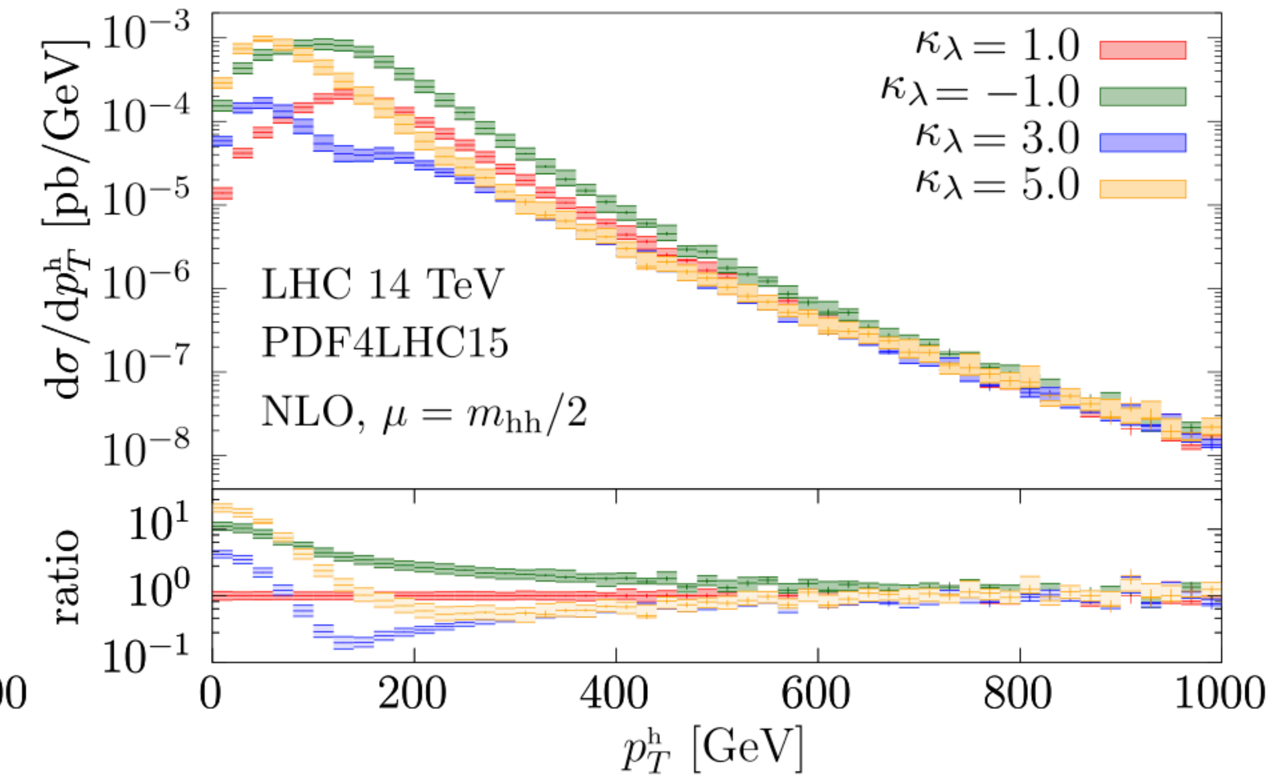
14/07/2023

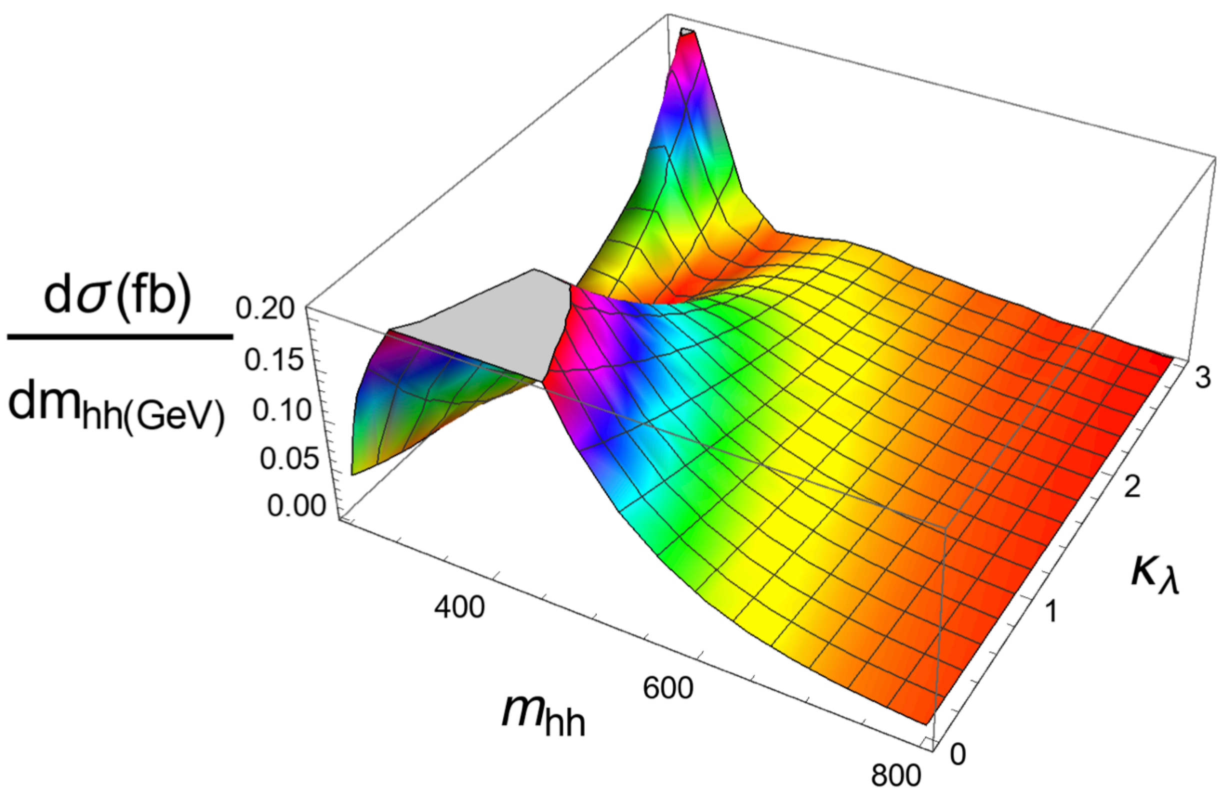
ICEPP
The University of Tokyo

UC San Diego

Di-Higgs production at higher energies







arXiv : 1910.00012

Constraining κ_λ is a major physics program of LHC and future colliders.

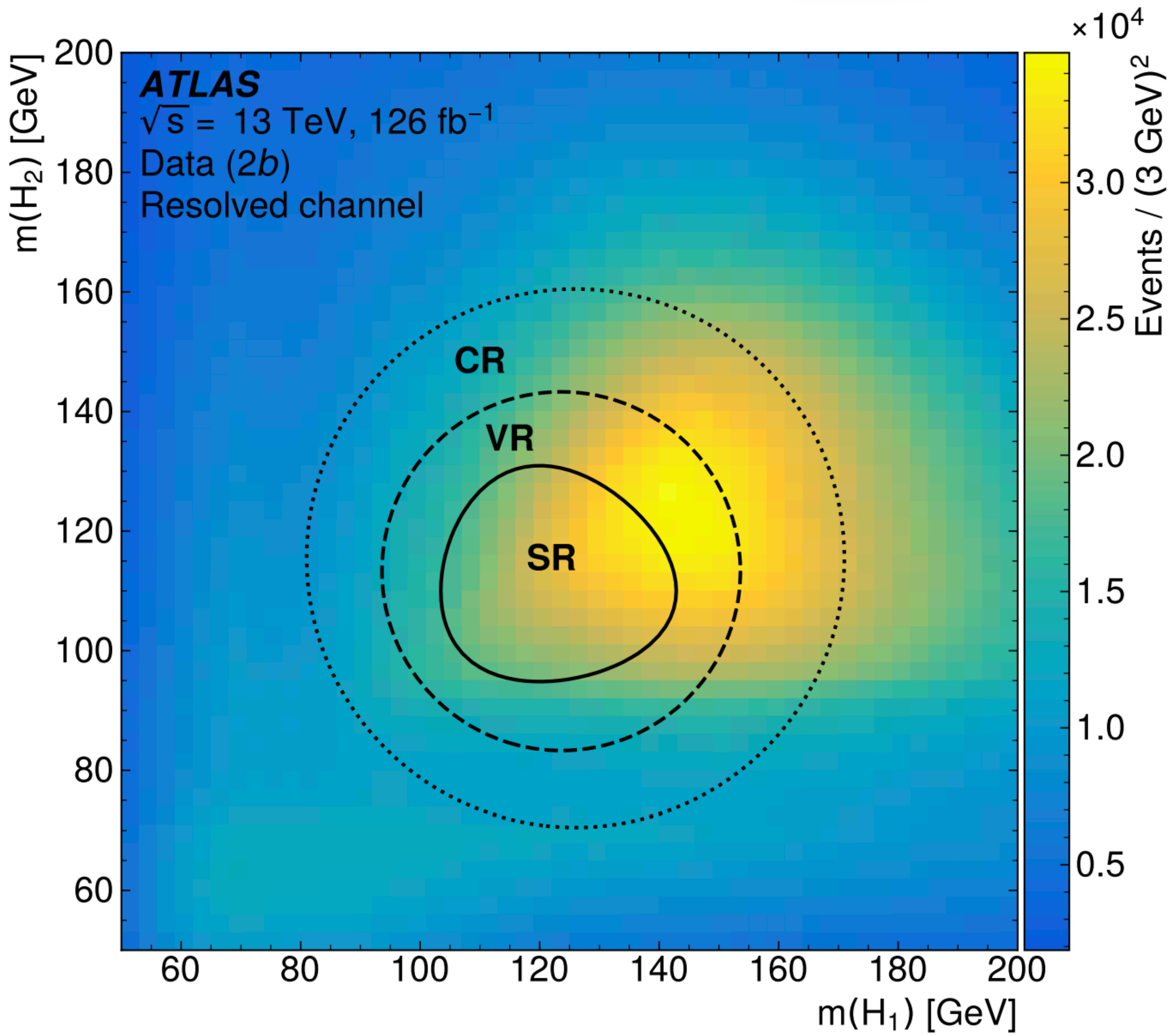
We aim to optimize a generic selection strategy for multi-Higgs final state in all-hadronic channel.

Resonant Higgs-pair : ATLAS

$$R_{HH}^{VR} \equiv \sqrt{(m(H_1) - 1.03 \times 120\,\text{GeV})^2 + (m(H_2) - 1.03 \times 110\,\text{GeV})^2} < 30\,\text{GeV}$$

$$X_{HH} = \sqrt{\left(\frac{m(H_1) - 120\,\text{GeV}}{0.1 \times m(H_1)}\right)^2 + \left(\frac{m(H_2) - 110\,\text{GeV}}{0.1 \times m(H_2)}\right)^2}$$

$$R_{HH}^{CR} \equiv \sqrt{(m(H_1) - 1.05 \times 120\,\text{GeV})^2 + (m(H_2) - 1.05 \times 110\,\text{GeV})^2} < 45\,\text{GeV}$$



$SR : X_{HH} < 1.6$

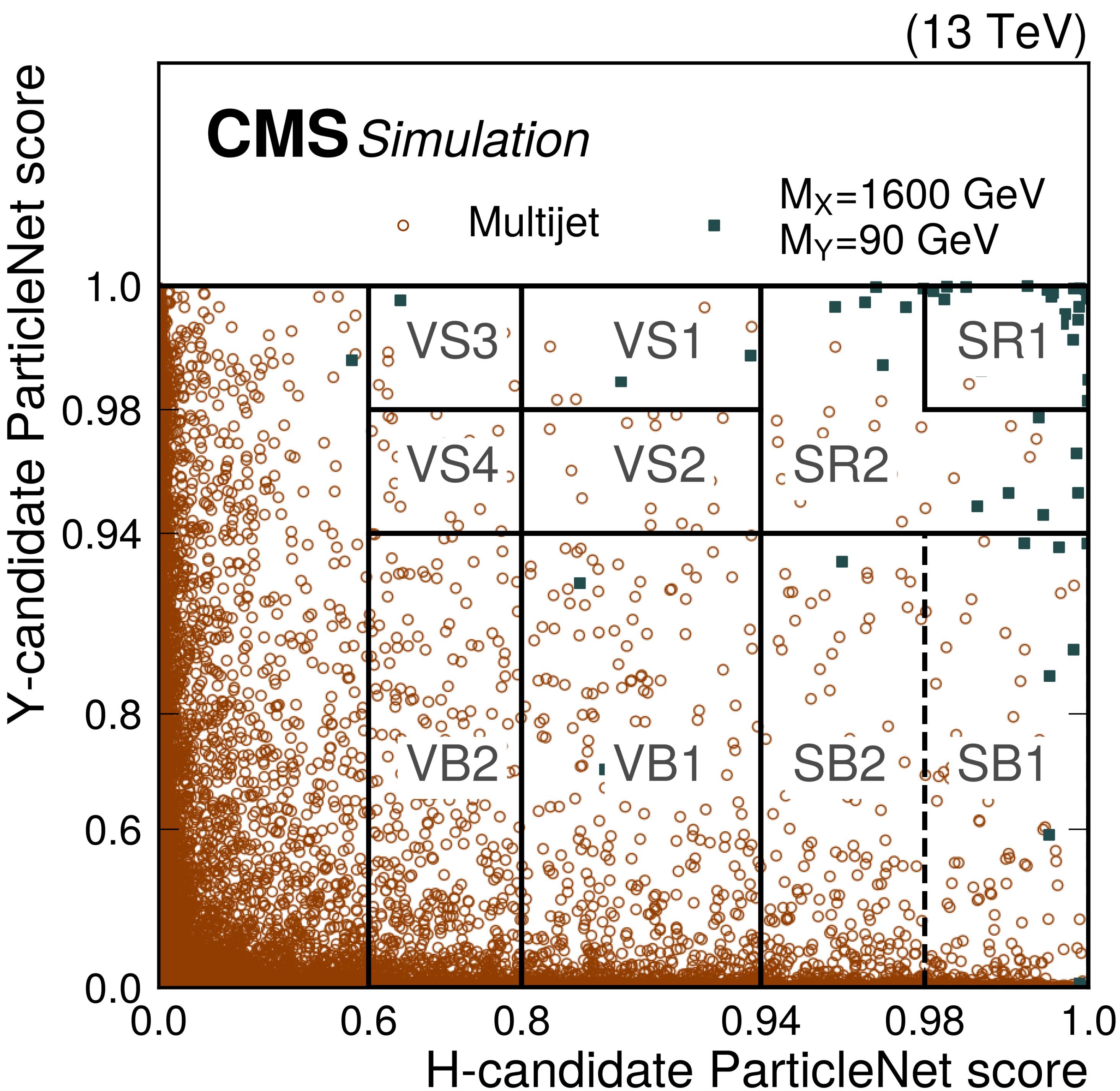
The pairing is chosen using a BDT trained by LightGBM (gradient BDT)

Phys. Rev. D 105 (2022) 092002

Y+H interpretations of 4b : CMS

$H \rightarrow b\bar{b}$ tagging is done using ParticleNet (GNN network)

Phys. Lett. B 842 (2023) 137392



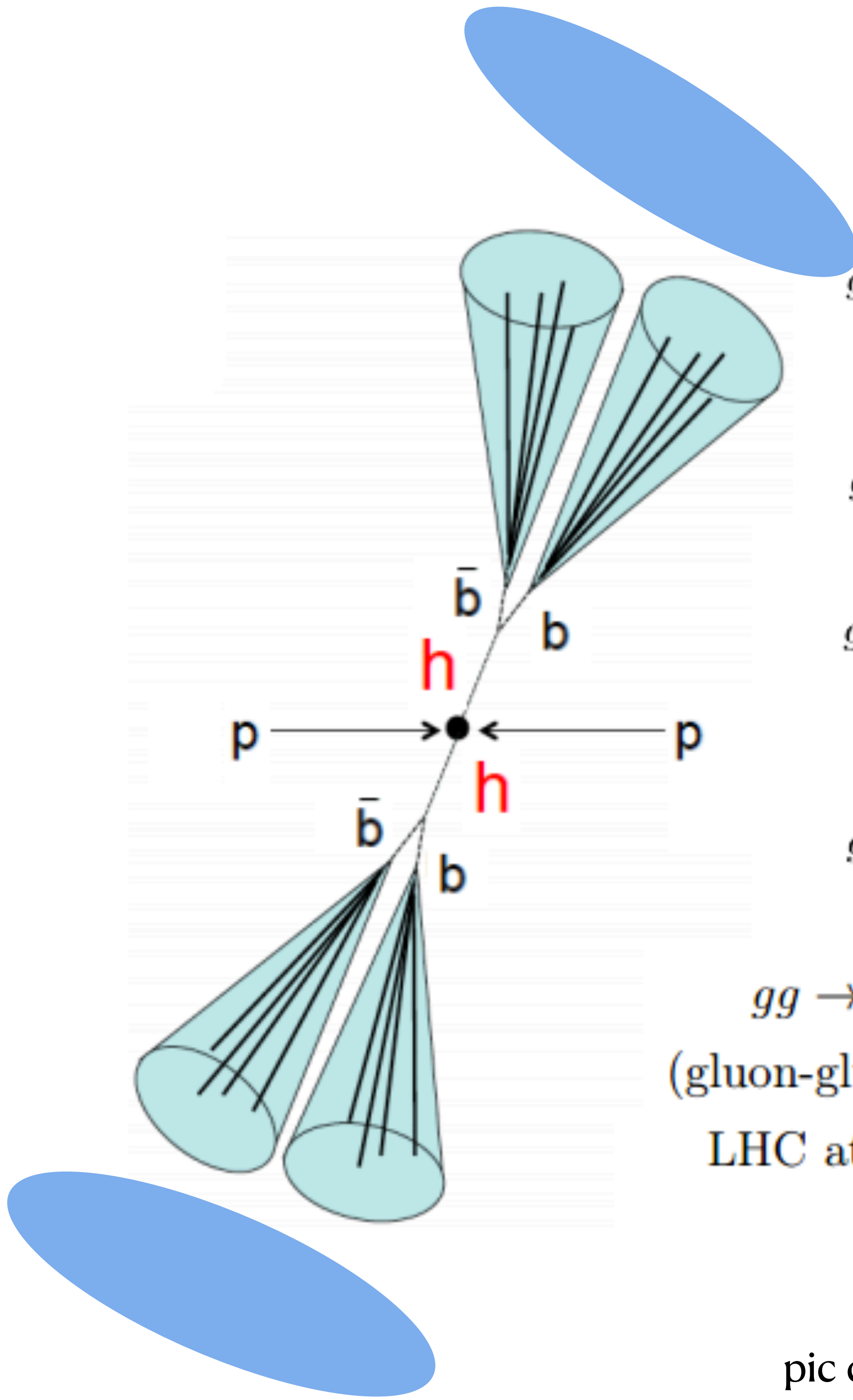
$$\boxed{P(\mathrm{R} \rightarrow \mathrm{bb})/(P(\mathrm{R} \rightarrow \mathrm{bb}) + P(\mathrm{QCD}))}$$

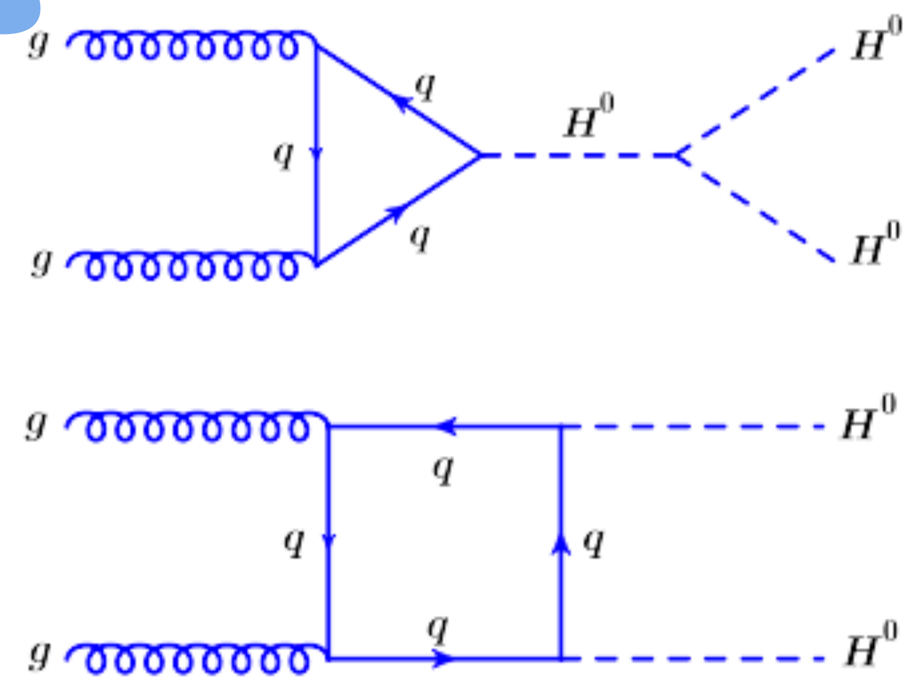
Same particlenet tagger was used for H->cc analysis.

Object identification using NN is now the standard prescription.

Event classification using GNN

For boosted di-Higgs production we look for two Ak-08 jets with track subjets





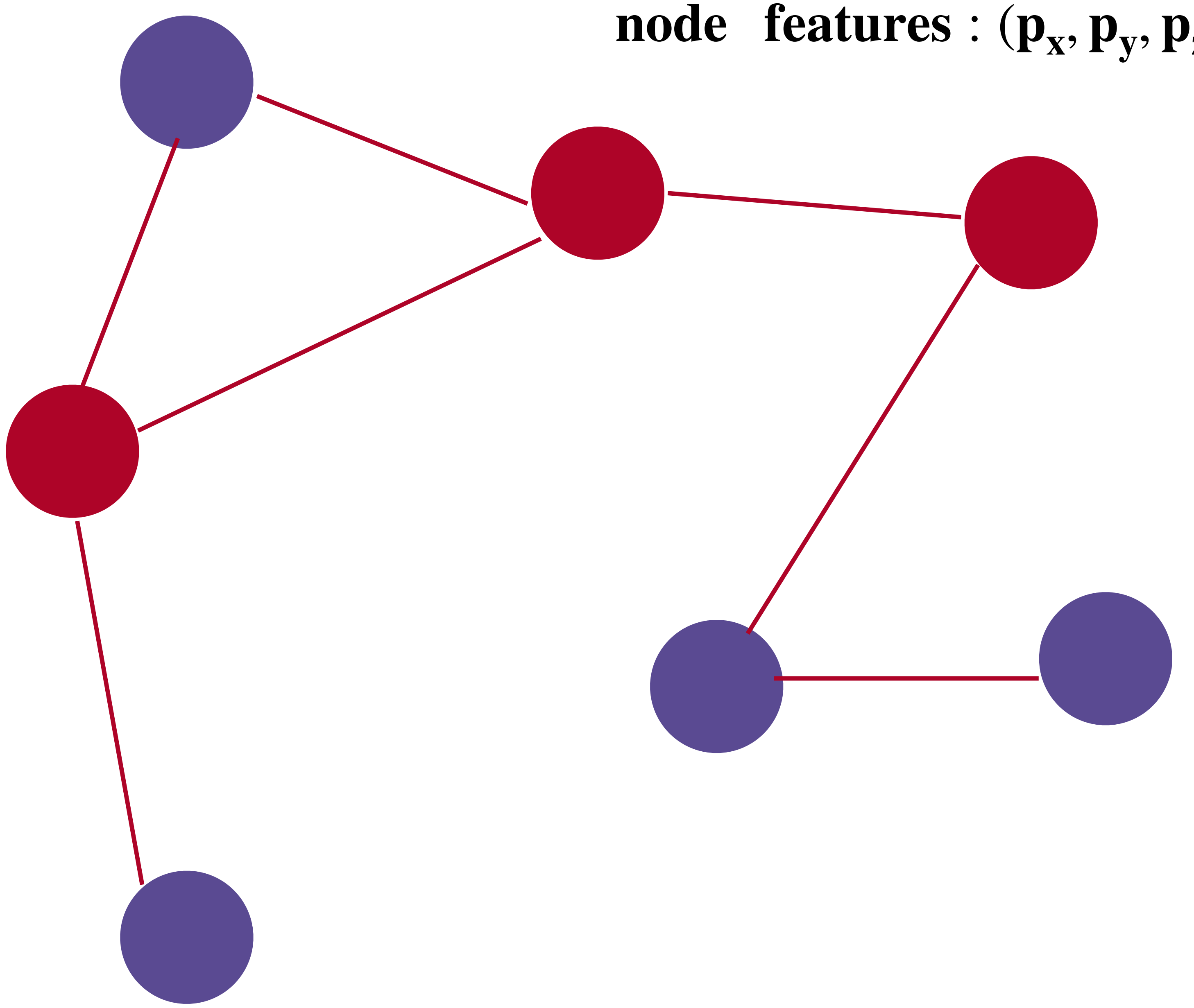
$gg \to H^0 H^0$
(gluon-gluon fusion)
LHC at $\sqrt{s} = 14$

$\sigma(pp \to H^0 H^0)$ [fb]	
NNLO	$40.2^{+3.2}_{-3.5}$

pic credit : https://francis.naukas.com/2014/08/22/el-campo-de-higgs/

Event as a graph

node features : (p_x, p_y, p_z, E, two_sub, one_sub)



Locate all the R=0.8 PF jets and 0.4 track jets in the η , ϕ plane.
Connect the k-NN neighbour through edges.
For each nodes : assign 4-vector + two and one subjettiness observables.
Use this graph representation for the events to be fed in GNN.

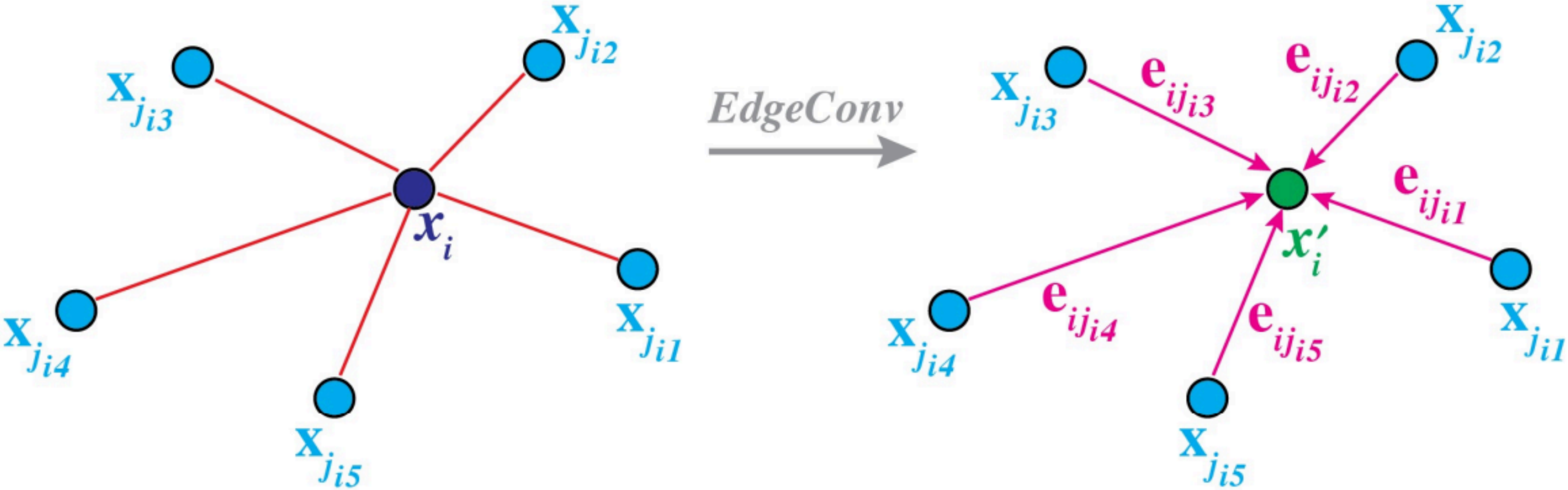
The graph network

https://arxiv.org/pdf/1801.07829.pdf

In a graph, each node can "learn" about the state of neighboring node through message passing operation

$$(x')_i^{l+1} = max_{j \in \mathcal{N}(i)} \Theta_x(x_j^l - x_i^l) + \Phi_x(x_i^l)$$

$$(e')_i^{l+1} = mean_{j \in \mathcal{N}(i)} \Theta_e(e_j^l - e_i^l) + \Phi_e(e_i^l)$$



The graph network

https://arxiv.org/pdf/1801.07829.pdf

$$(x')_i^{l+1} = max_{j \in \mathcal{N}(i)} \, \Theta_x(x_j^l - x_i^l) + \Phi_x(x_i^l)$$

$$(e')_i^{l+1} = mean_{j \in \mathcal{N}(i)} \, \Theta_e(e_j^l - e_i^l) + \Phi_e(e_i^l)$$

Learned "coordinates"

$(\overrightarrow{x_1}, E_1)$ (z_1^M, E_1^{L1})

$(\overrightarrow{x_2}, E_2)$ (z_2^M, E_2^{L1})

$(\overrightarrow{x_3}, E_3)$ $\longrightarrow$ $\mathbf{EC_{L1}(d_{in} = 3,\ d_{out} = M)}$ $\longrightarrow$ (z_3^M, E_3^{L1})

Learned "energy"

.
.
.

$(\overrightarrow{x_N}, E_N)$ (z_N^M, E_N^{L1})

**Forms a K-NN graph
and performs
the message passing**

The graph network

https://arxiv.org/pdf/1801.07829.pdf

$(\overrightarrow{x_1}, E_1)$

$(\overrightarrow{x_2}, E_2)$

$(\overrightarrow{x_3}, E_3)$

.

.

.

$(\overrightarrow{x_N}, E_N)$

$\longrightarrow$ $\mathbf{EC_{L1}(d_{in} = 3,\ d_{out} = M)}$ $\longrightarrow$

Forms a K-NN graph
and performs
the message passing

(z_1^M, E_1^{L1})

(z_2^M, E_2^{L1})

(z_3^M, E_3^{L1})

.

.

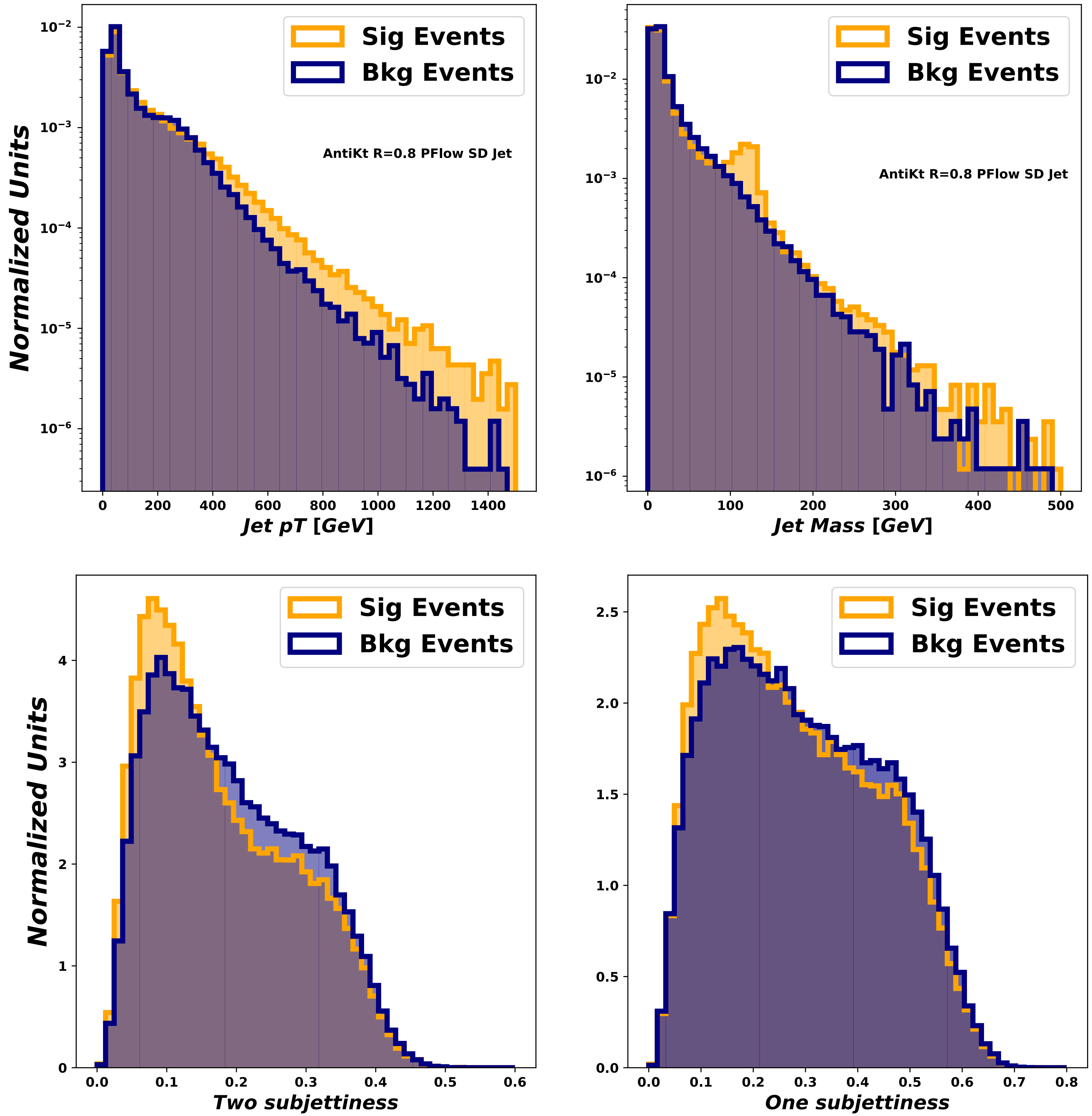
.

(z_N^M, E_N^{L1})

After p message passing layers , the q-th node has following energy representation :

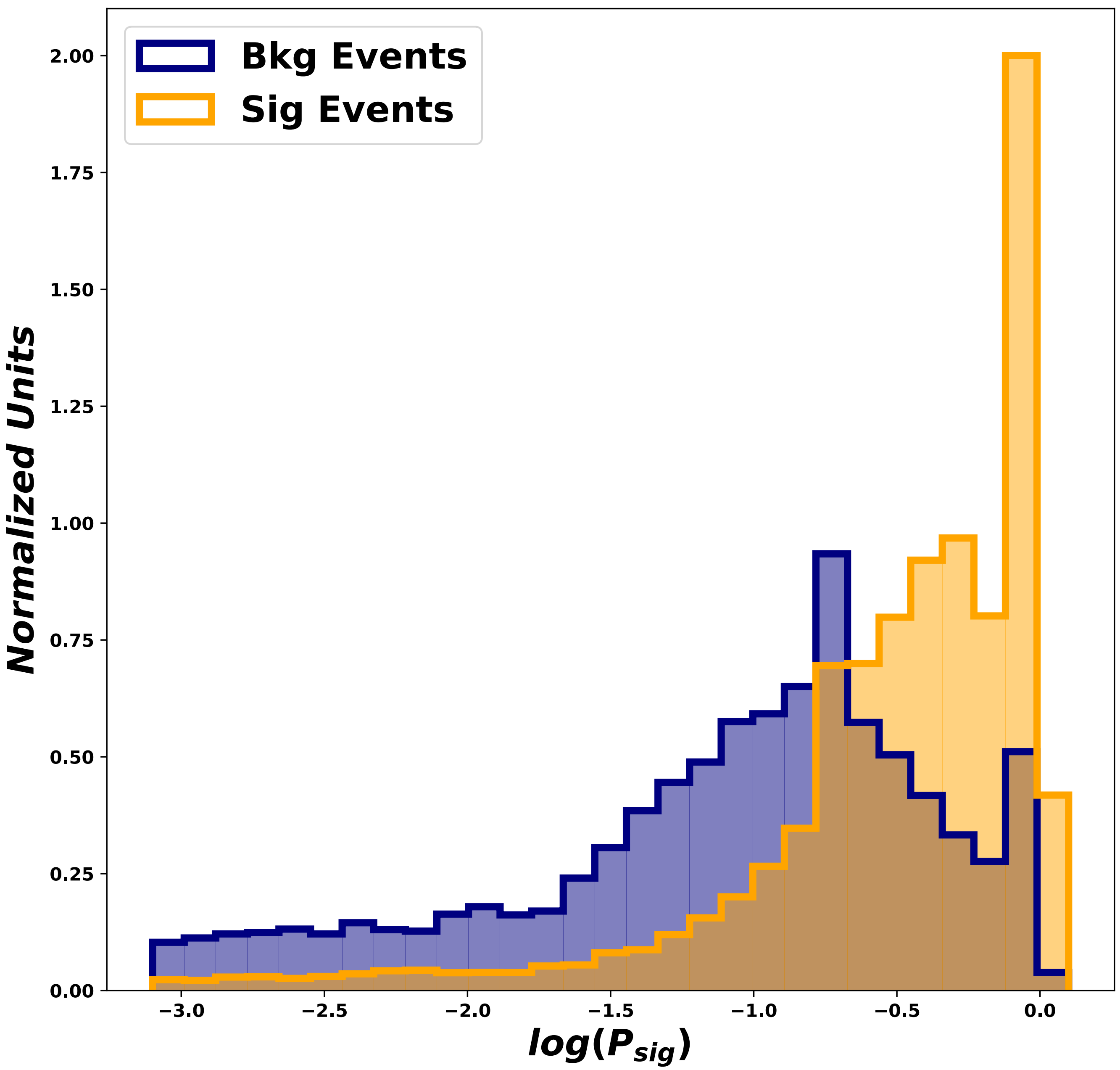
$$\left(\mathbf{E_q}, \mathbf{E_q^{L1}}, \mathbf{E_q^{L2}}, \ldots, \mathbf{E_q^{Lp}} \right) \longrightarrow \textbf{MLP} \longrightarrow [p_{sig}, p_{bkg}]$$

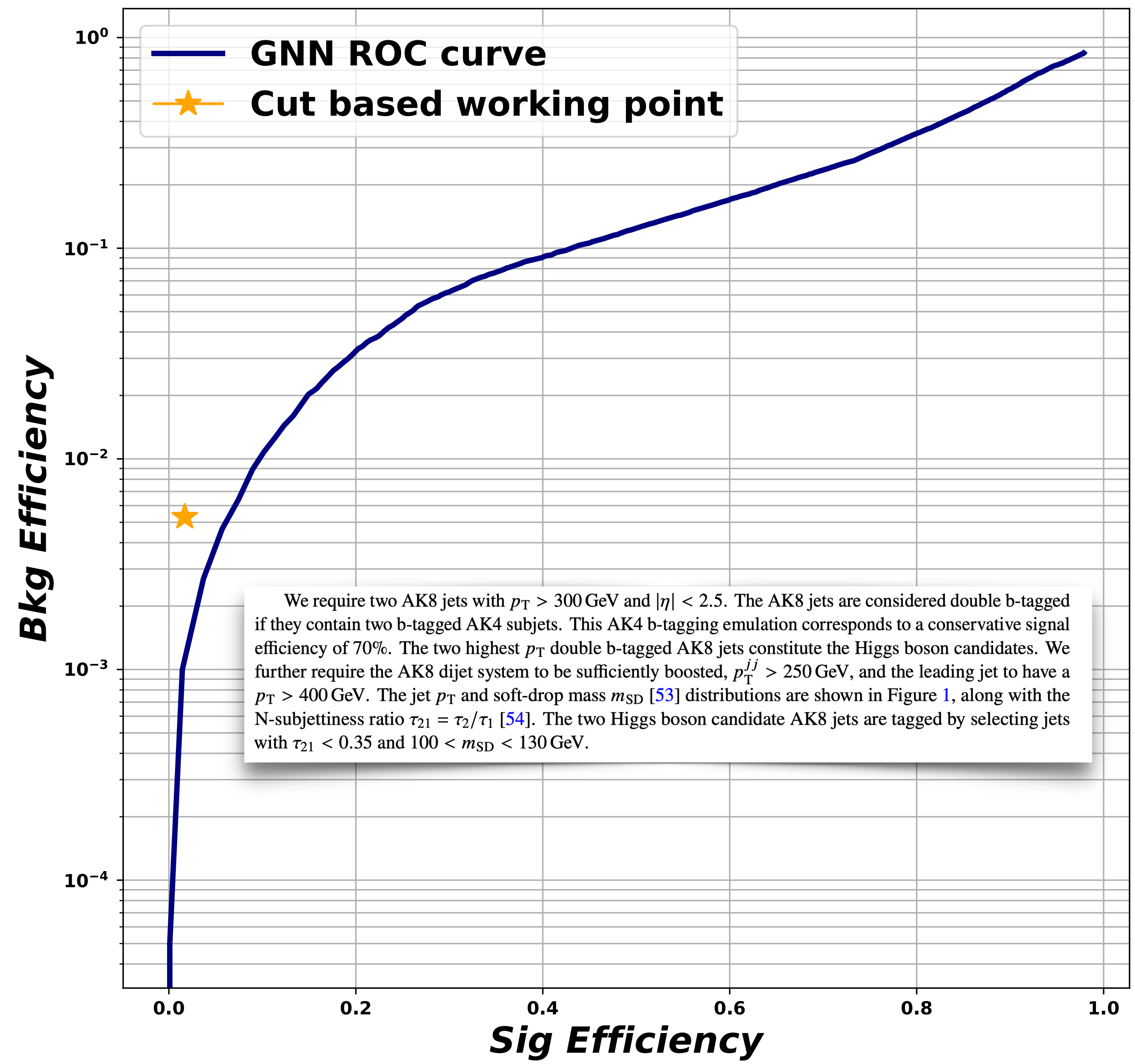
Some distributions of the features



arXiv:2203.07353

Training outputs





We require two AK8 jets with $p_T > 300\,\text{GeV}$ and $|\eta| < 2.5$. The AK8 jets are considered double b-tagged if they contain two b-tagged AK4 subjets. This AK4 b-tagging emulation corresponds to a conservative signal efficiency of 70%. The two highest p_T double b-tagged AK8 jets constitute the Higgs boson candidates. We further require the AK8 dijet system to be sufficiently boosted, $p_T^{jj} > 250\,\text{GeV}$, and the leading jet to have a $p_T > 400\,\text{GeV}$. The jet p_T and soft-drop mass m_{SD} [53] distributions are shown in Figure 1, along with the N-subjettiness ratio $\tau_{21} = \tau_2/\tau_1$ [54]. The two Higgs boson candidate AK8 jets are tagged by selecting jets with $\tau_{21} < 0.35$ and $100 < m_{SD} < 130\,\text{GeV}$.

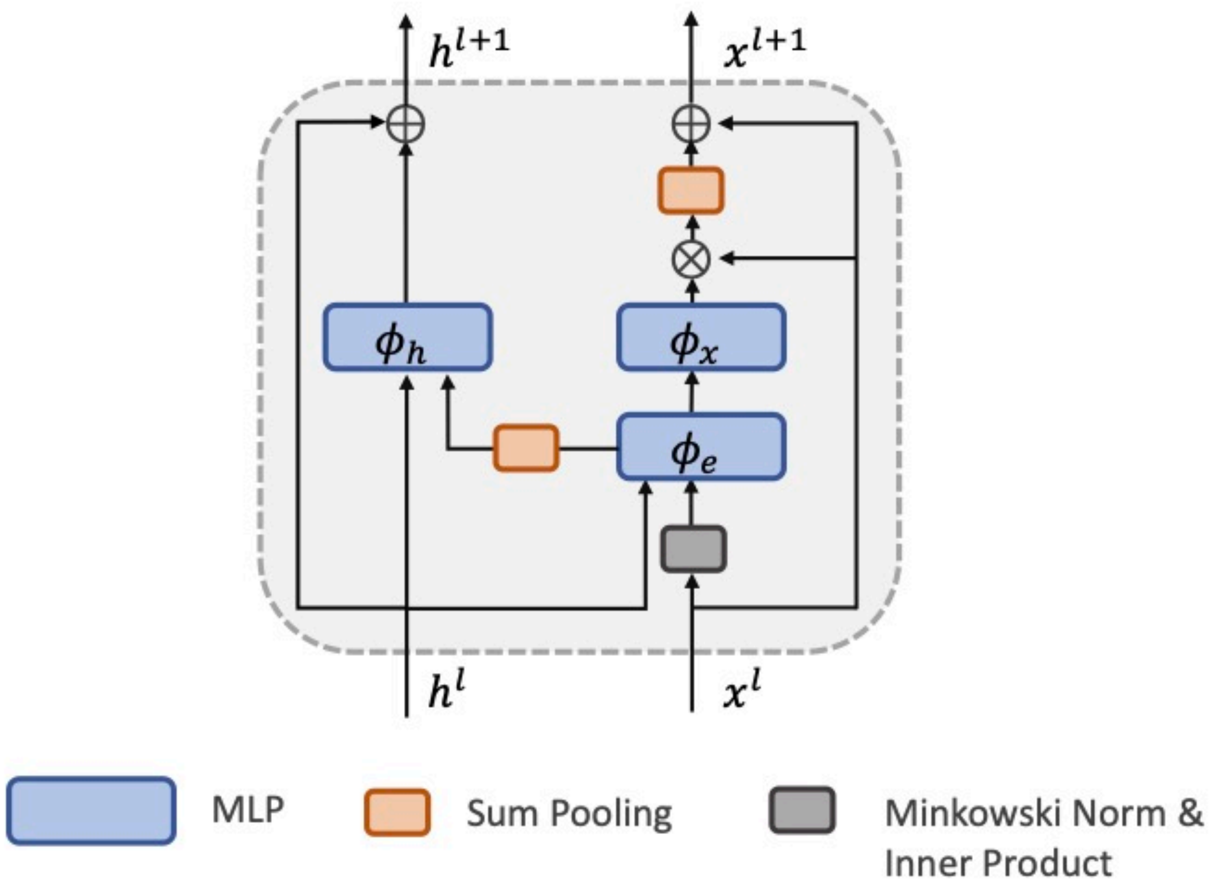
The preliminary NN has 3X times background rejection for the same efficiency around cut-based WP

An attempt with LE GNN

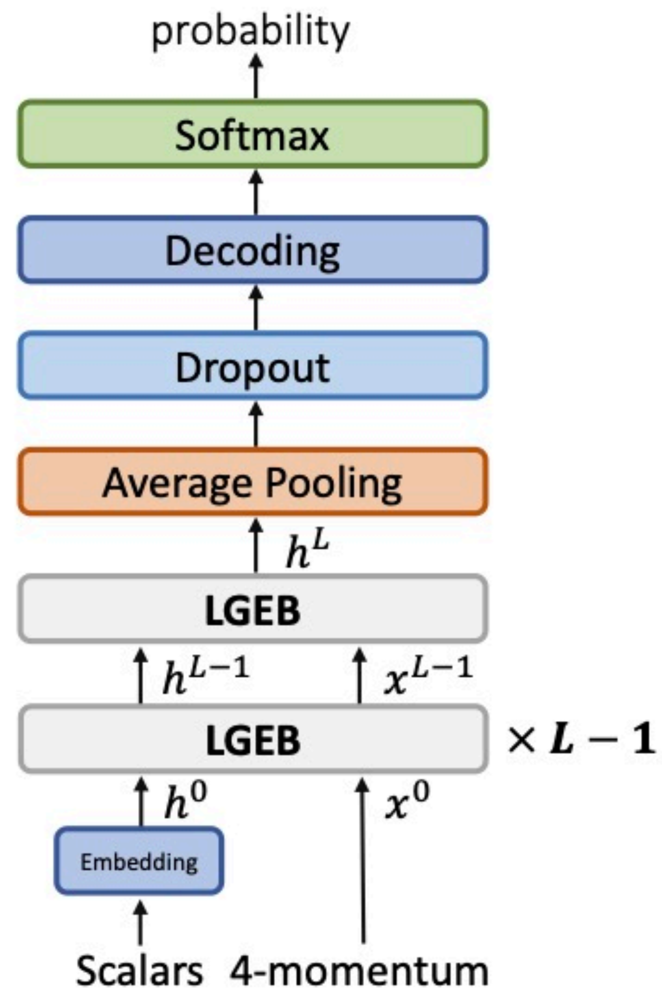
https://arxiv.org/pdf/2201.08187.pdf

$$m_{ij}^l = \phi_e\left(h_i^l, h_j^l, \psi(\|x_i^l - x_j^l\|^2), \psi(\langle x_i^l, x_j^l\rangle)\right)$$

$$x_i^{l+1} = x_i^l + c\sum_{j\in[N]}\phi_x(m_{ij}^l)\cdot x_j^l \qquad h_i^{l+1} = h_i^l + \phi_h(h_i^l, \sum_{j\in[N]} w_{ij}m_{ij}^l)$$

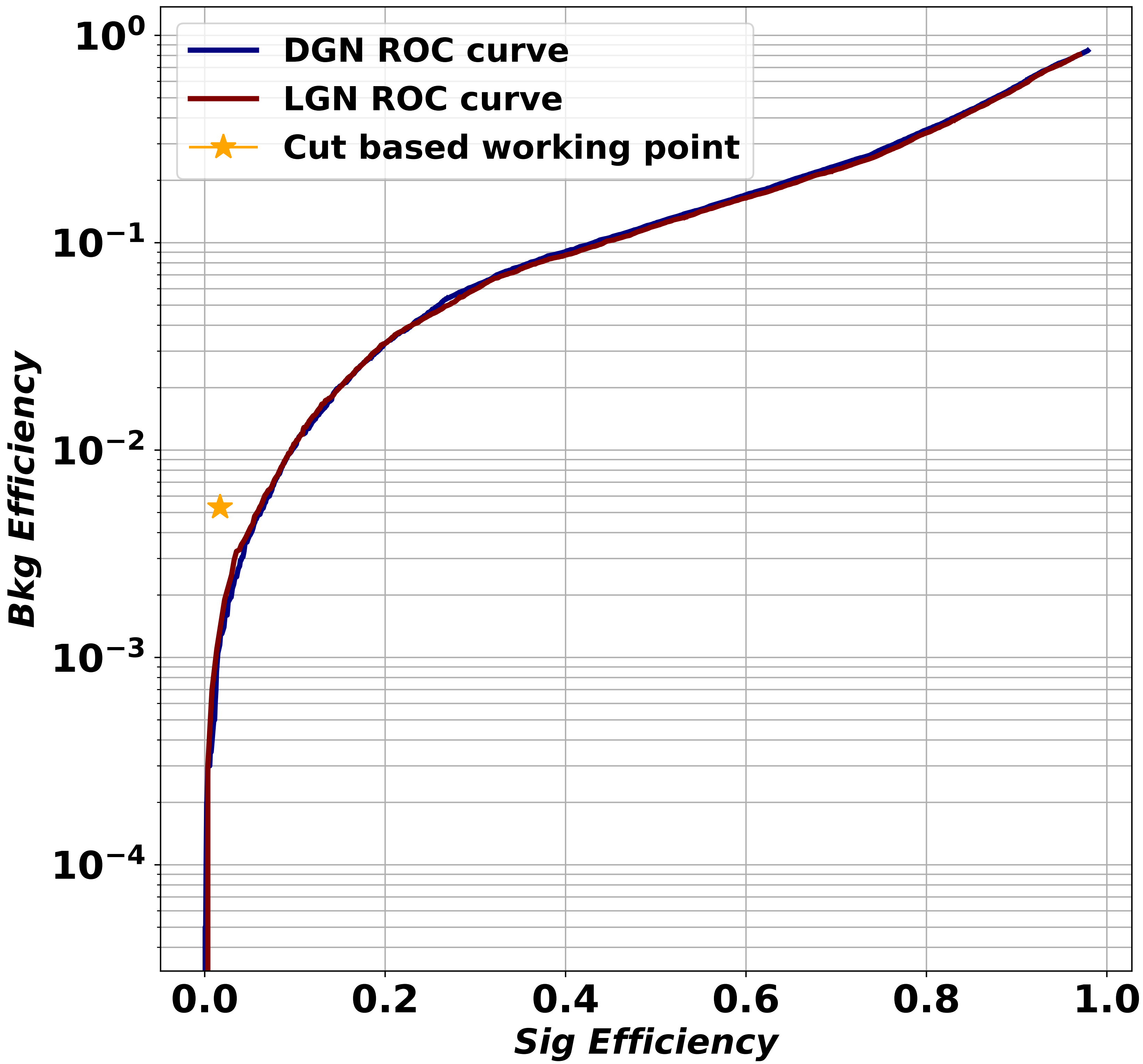


Lorentz Group Equivariant Block (LGEB)



LorentzNet

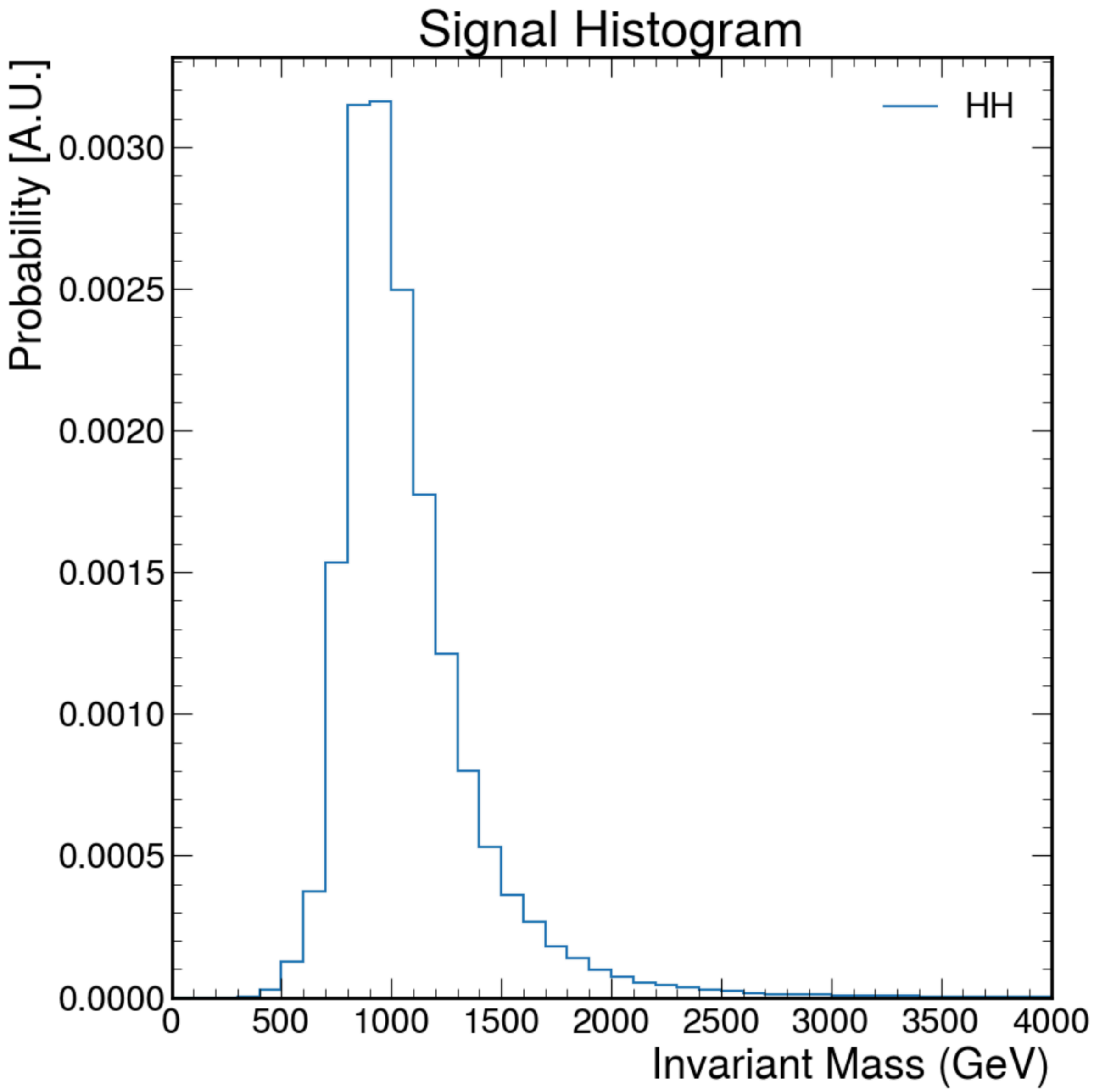
ROC curves comparison

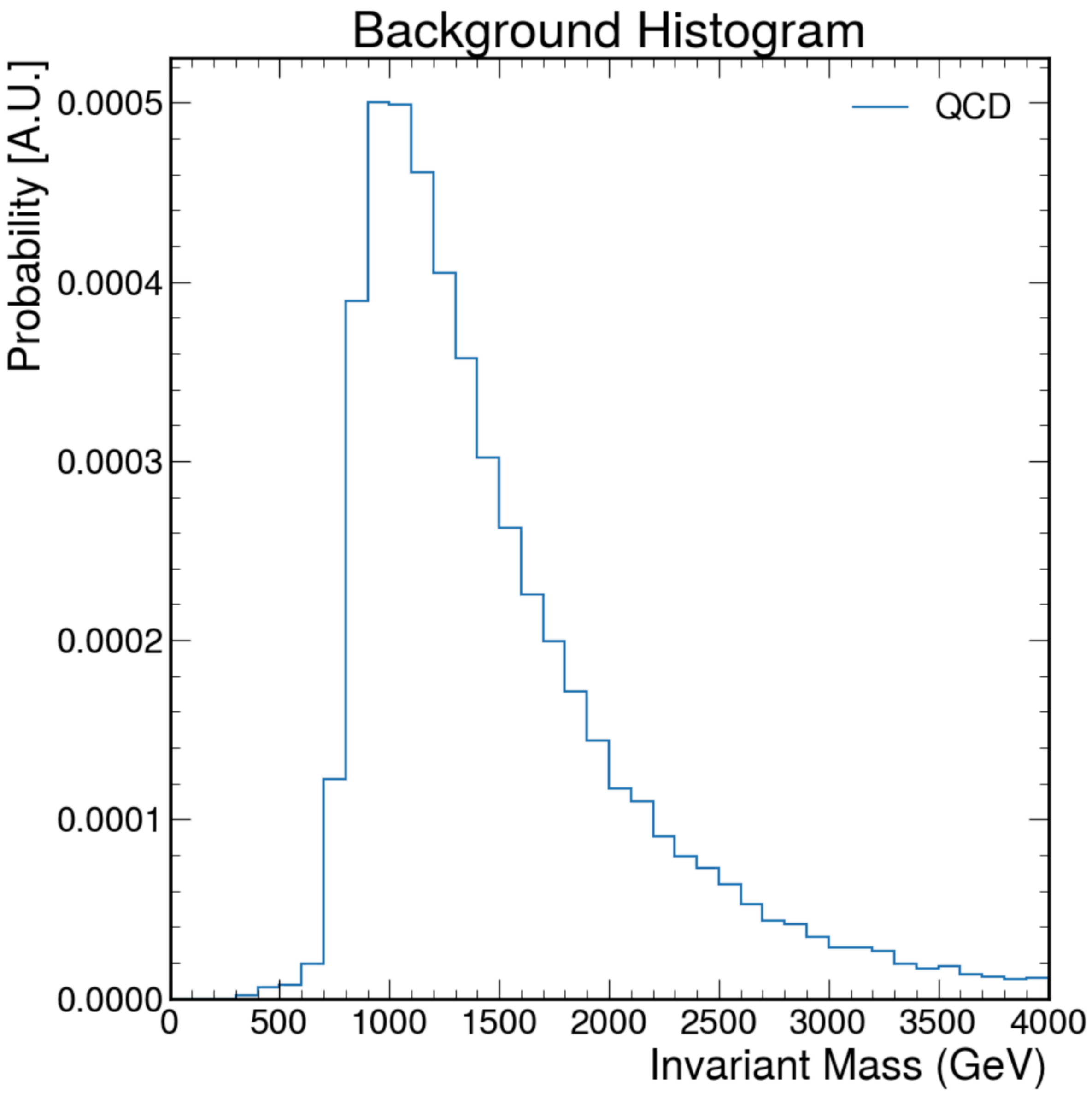


Both the MP-networks have similar event classification efficiency.

Pre & post selection distributions

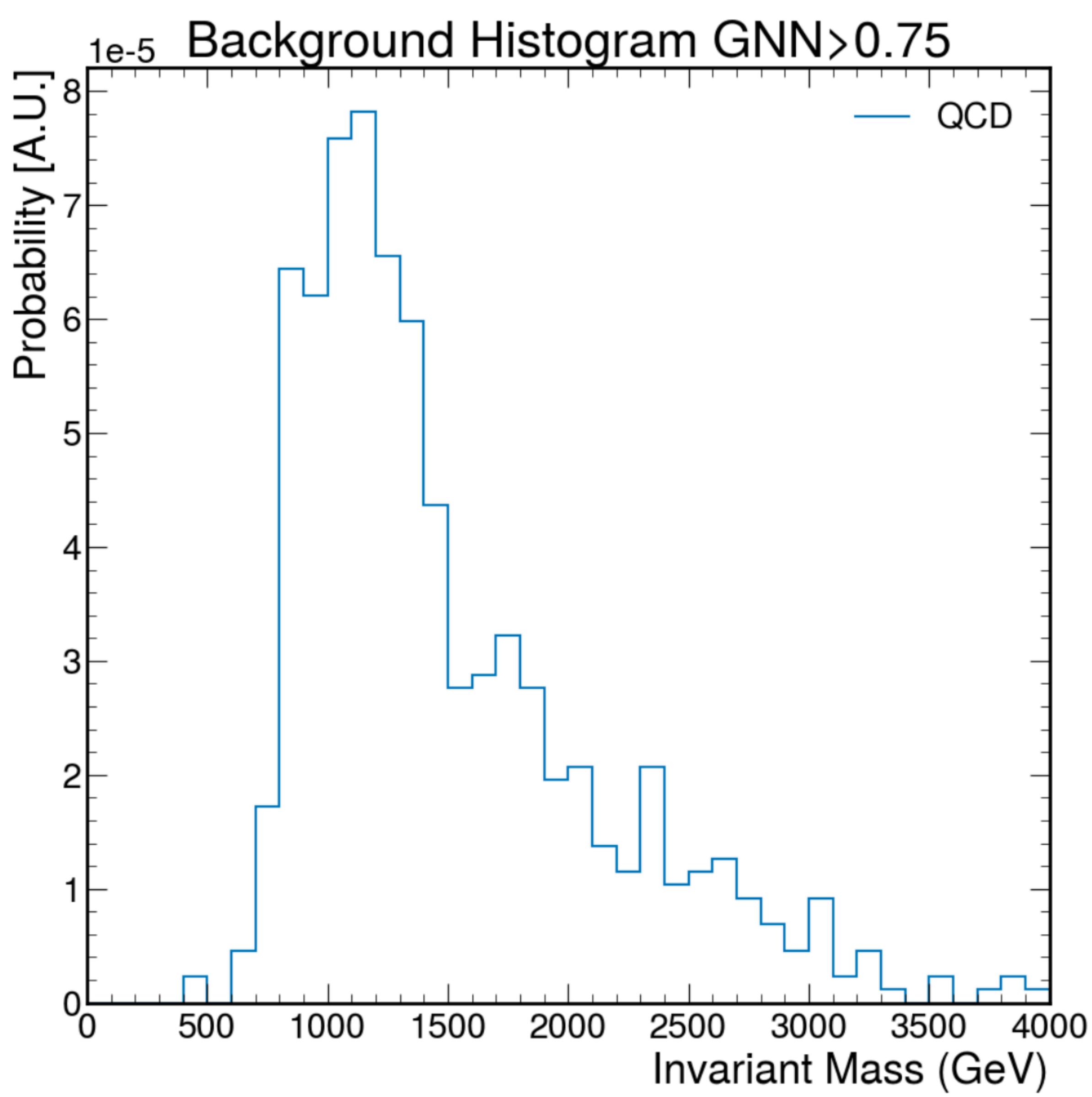
- ☑ Large-R jet having two subjets.
- ☑ Jet1 pT > 400 GeV, Jet2 pT > 300 GeV
- ☑ |Jet eta| < 2.5

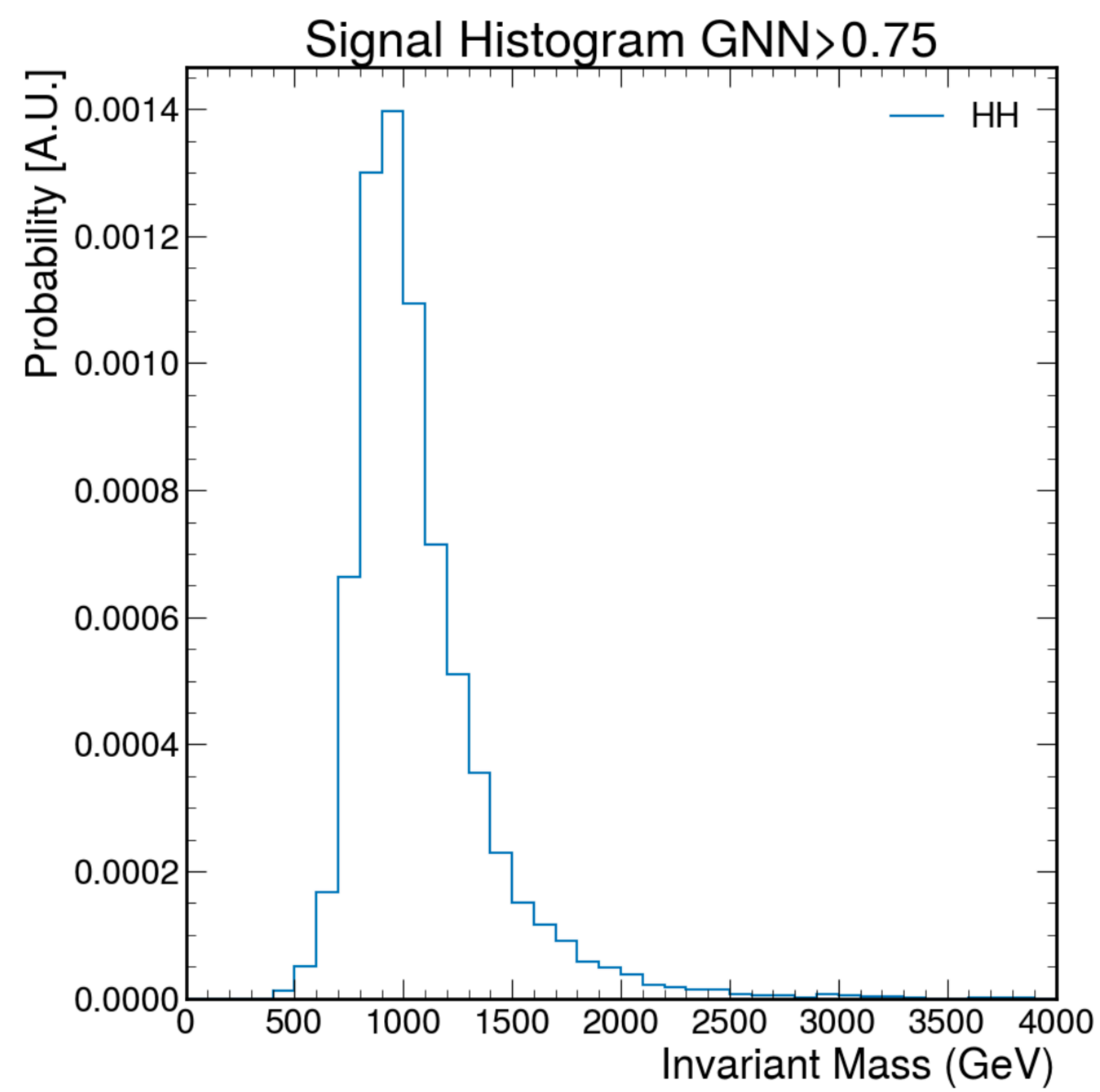




Pre & post selection distributions

- **Large-R jet having two subjets.**
- **Jet1 pT > 400 GeV, Jet2 pT > 300 GeV**
- **|Jet eta| < 2.5**

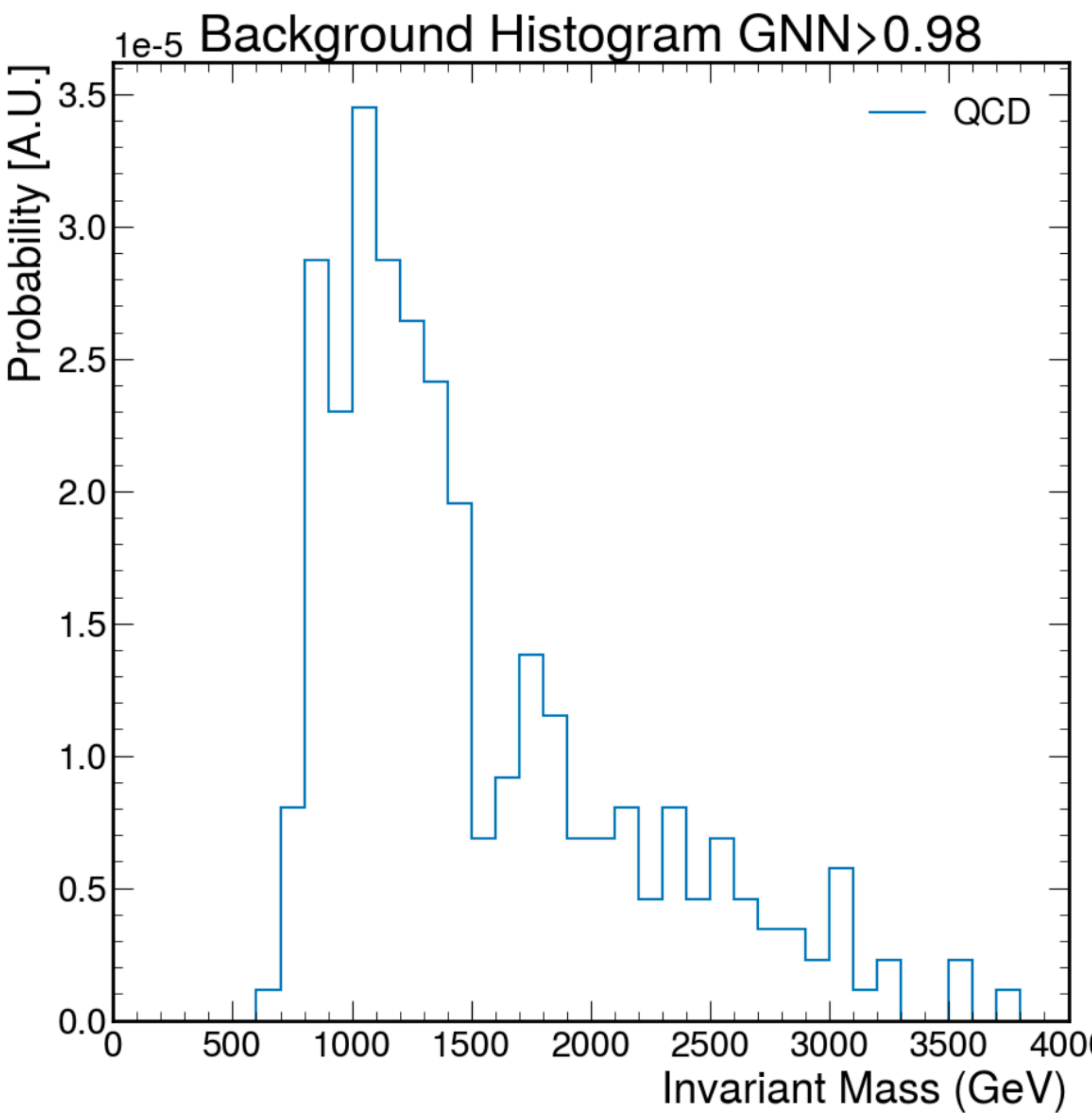


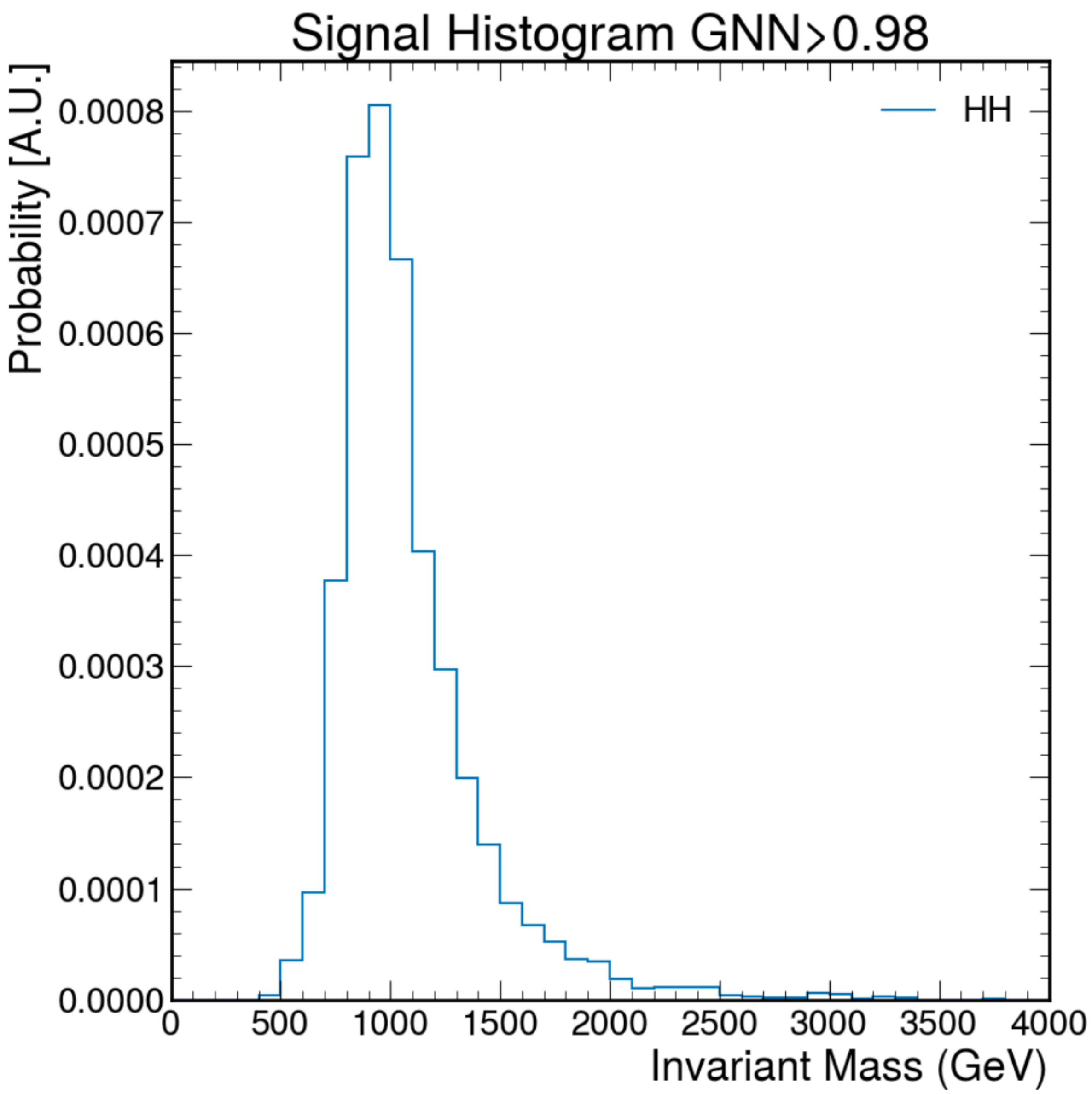


Pre & post selection distributions

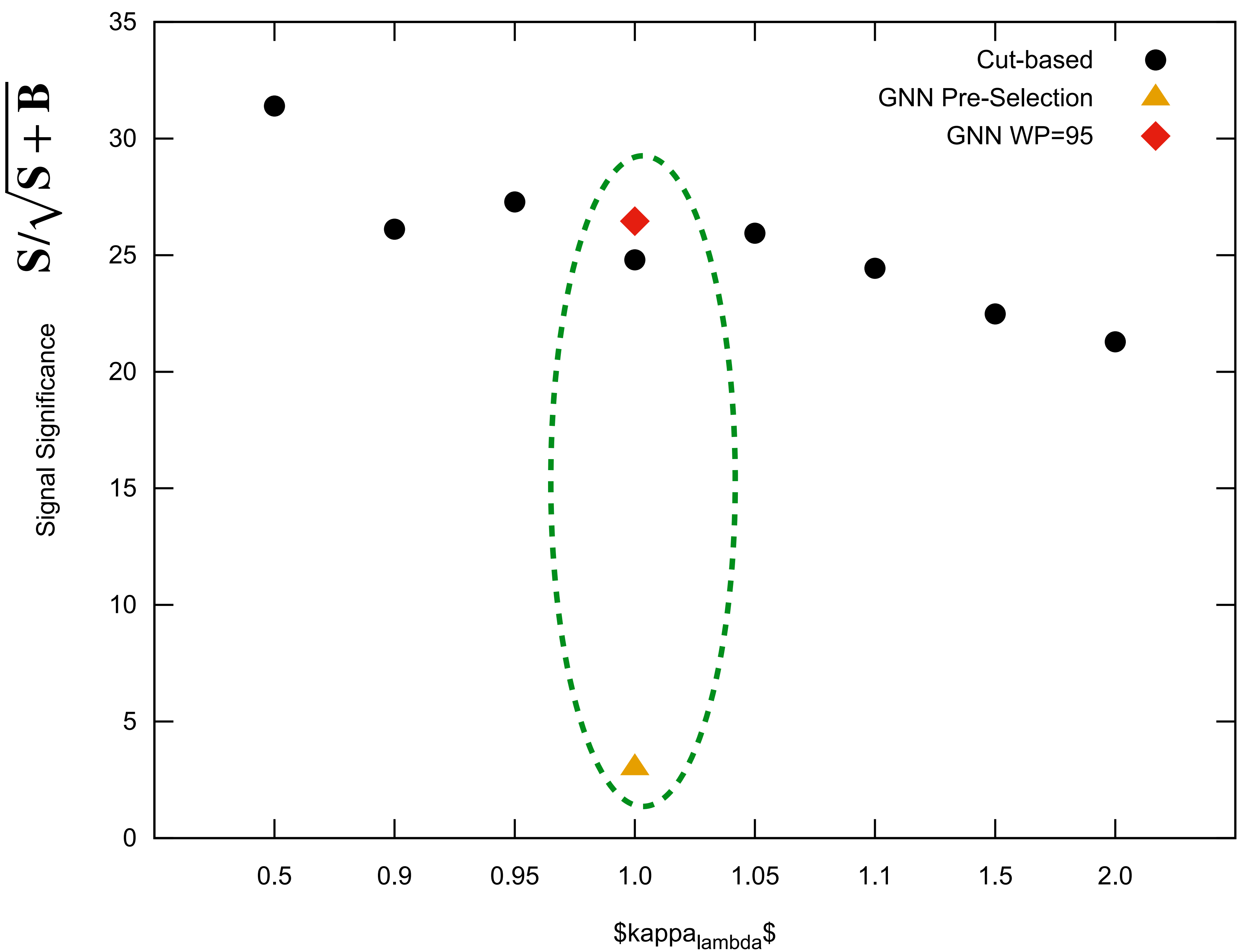
- ☑ Large-R jet having two subjets.
- ☑ Jet1 pT > 400 GeV, Jet2 pT > 300 GeV
- ☑ |Jet eta| < 2.5

Larger stat fluctuation on QCD tail due to rejected events.





Pre & post selection distributions



Take away

- ☑ For multi-b final states event level classifiers are capable of increasing sensitivity
- ☑ These preliminary studies (for FCC-hh) are generalizable across hh or hhh searches
- ☑ Compared two different GNN models : probably a general GNN will do the required job.
- ☑ The individual tagging score and pairing scores should improve the sensitivity.

THANK YOU !!