ML Particle Assignment in ttH Decay

Christian B. Hughes

Czech Technical University in Prague

Prague, February 2026

Outline

The ideas in these slides constitute an overview of Adam Herold's FEL ČVUT thesis. The figures are sources from the same thesis.

- What is ttH
- Detector measurements
- Particle assignment
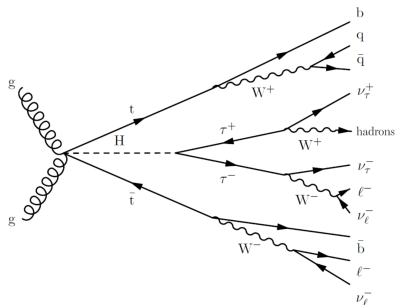- Particle assignment ML pipeline
- Outputs

What is ttH ? (recap)

- For gluons g, we sometimes produce :

$$gg \to t\bar{t}H$$

- The Higgs H decays into lighter particles.
- The top quarks $t, \bar{t}$ can decay as :

$$t \to Wb$$

- The boson W can then decay either to $\ell\nu$ or to $q\bar{q}$.

What the detector gives us

The detector does not observe quarks/top/Higgs directly. We reconstruct from leptons, jets, and missing tranverse energy.

The specific analysis channel

A common selection is :

$$2\ell SS + 1\tau_{\mathrm{had}}$$

for leptons ℓ. We have

- Two leptons with the same charge.

- One hadronic tau.

- Multiple jets.

The event is consistent with a ttH-like decay pattern, but the mapping from observed objects to decayed is ambiguous.

Combinatorics

Suppose we have N_j jets and a decay that requires us to decide "this jet is b from top" and "that jet is from W," and so on.

Then the number of plausible choices (assignments) is :

$$\# \text{ assignments} = \#\text{choices} \times \#\text{orderings}.$$

If incorrect assignment occurs, everything downstream of it (including Higgs mass reconstruction) is unreliable.

Additional geometric information

We measure directions using (η, ϕ), where :

- ϕ : azimuth angle around the beam axis,
- η : pseudorapidity (an angle coordinate in collider geometry).

The distance between reconstructed objects is :

$$\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}.$$

This allows us to discern whether assignments are plausible or implausible.
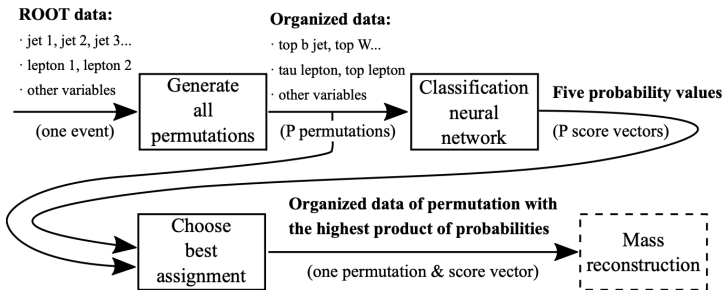
The goal of particle assignment

Given a decay pattern, which reconstructed objects fill which roles in the decay ?
By roles, we mean things like : "b from top" or "jet from W. Observations are specific jets/leptons/τ_{had}.
To achieve this, we enumerate candidate assignments, score them, and select the "best" according to some metric of success.

Assignment use in the reconstruction pipeline (high level)

- **Step 1** : Generate all permutations of objects into the required roles.
- **Step 2** : For each permutation, compute physics-motivated features.
- **Step 3** : A neural network outputs plausibility scores.
- **Step 4** : Choose the assignment with the best overall score.

Candidate generation

Suppose we have a set of jets $\{j_1, \ldots, j_{N_j}\}$. The decay pattern requires selecting certain jets into distinct roles. The number of permutations grows rapidly. We apply basic selection rules to keep the number of permutations manageable.

What makes a "good" assignment ? (intuition)

A correct assignment tends to make :

- angular separations ΔR plausible.
- *b*-tagged jets often land in *b*-quark roles
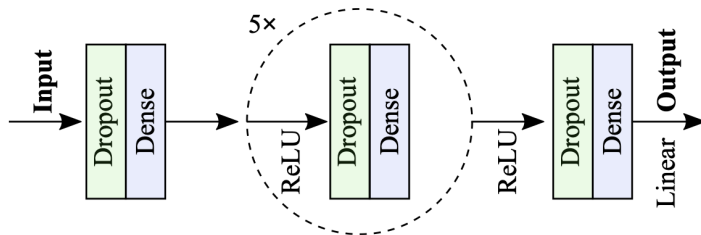- overall event kinematics look consistent with the decay pattern.

The neural network learns these tendencies from labeled simulation data (also called the truth information).

Inputs : what features does the NN see ?

We feed a fixed-length vector of physics + detector features (example categories) :

- Object kinematics :

    - p_T, η, ϕ, m

- Geometry :

    - ΔR between candidate objects

- $\Delta\phi$, $\Delta\eta$

- Masses :

    - (e.g., $m(jj)$, $m(\ell j)$)

- Tagging :

    - b-tags

Architecture (one network, multiple role-scores)

- Shared hidden layers learn a representation of candidate assignment.
- Output heads give per-role probabilities.

Final assignment selection

For each candidate assignment a, the NN returns role-scores :

$$p_1(a), p_2(a), \ldots, p_K(a)$$

We choose :

$$a^{\star} = \arg \max_{a \in \mathcal{A}} \prod_{k=1}^{K} p_k(a).$$

That is, we select the candidate whose full set of role decisions is most self-consistent.

Why assignment helps Higgs mass reconstruction

Later tasks (mass regression/separation) assume the right objects are used. Incorrect assignment causes error propagation and creates noise.

Summary

- Observe jets/leptons/τ.
- Enumerate candidate assignments and score each with a NN using physics-informed features.
- Choose the assignment maximizing a probabilistic consistency score.