A Precision Side Quest at FCC-ee: Probing the Electron–Yukawa Coupling

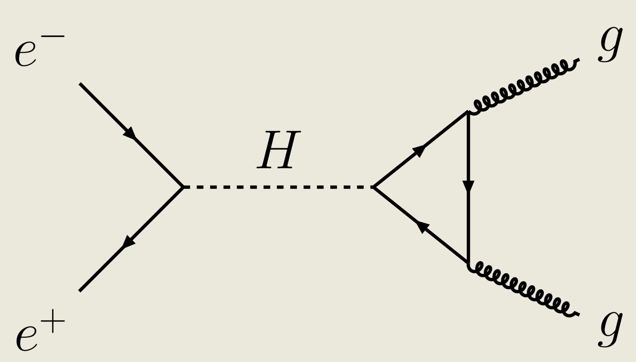
Hugo Leigh-Watts

Supervisor: Prof. Victoria Martin

University of Edinburgh

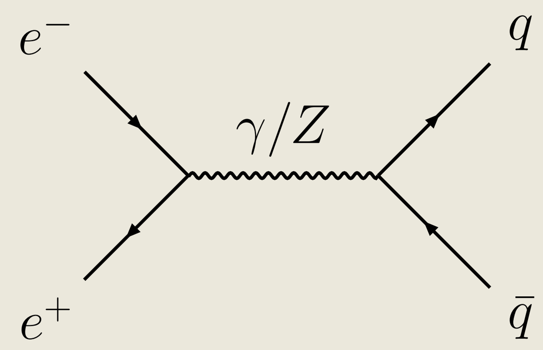
I. The Side Quest

Objective: Investigate whether quark/gluon jet-level observables can be distinguished to establish the feasibility of measuring the electron–Yukawa coupling at the Future Circular Collider (FCC).



Signal: $e^+e^- \to H \to gg$

from



Background: $e^+e^- \to \gamma/Z \to q\bar{q}$

Why $H \to gg$? [1]

Signal	
Branching fraction	$\mathcal{B}(H \to gg) = 8.2\%$
Cross section	$\sigma_{sig} = 23\,\text{ab}$
Background	
Process	$e^+e^- \to q\bar{q}$ (irreducible)
Cross section	$\sigma_{bkg} = 61\,\text{pb}$
Result	
Naive S/B	4×10^{-7}

Key Challenge: Need to distinguish between Gluon and quark jets!

Other decays have similarly large branching ratios but too much background ($H \to b\bar{b}$) or have exceptionally clean final states but small branching ratio ($H \to \gamma\gamma$) making them statistically negligible.

III. Methodology: Tools of the Hero

Analysis Pipeline

Event Generation (by CERN) $\to$ Feature Engineering $\to$ ML Classification

Simulated Data

- Using simulated event samples from the FCC-ee Winter 2023 Monte Carlo campaign (CERN) [3]
- Full detector simulation and jet reconstruction

	Signal $e^+e^- \to H \to gg$	**Background** $e^+e^- \to q\bar{q}$
Events simulated	1.2M	499M
Events CME	125GeV	125GeV
Detector	IDEA	IDEA
S/B	2×10^{-7}	

- Matches the order of magnitude from Ref. [1]

Feature Engineering:

- Can eliminate energies not around the Higgs mass
- Number of particles in a jet (Gluons have more particles)
- Jet width (Gluons are broader)

After feature engineering, we are left with signal and background that are very similar; this is where machine learning is useful to try and distinguish between jets.

II. Background: The World of FCC-ee

The FCC-ee is a proposed $\sim$91 km e^+e^- collider at CERN (2040s), running at multiple centre-of-mass energies (CME), one being at the **Higgs mass** (125GeV) for direct production.

Why FCC-ee, not LHC? At the LHC, measuring the Yukawa coupling, y_e, would require detecting $H \to e^+e^-$ with $\mathcal{B} = 5\times10^{-9}\,\%$ against overwhelming Drell-Yan backgrounds — effectively impossible.[1]

At FCC-ee ($\sqrt{s} = m_H$), we would use **s-channel production** $e^+e^- \to H$:

$$\sigma_{ee\to H} = \frac{4\pi\Gamma_H\Gamma(H \to e^+e^-)}{(s - m_H^2)^2 + m_H^2\Gamma_H^2}$$

The advantage: We can use large branching ratio decays ($H \to b\bar{b}$, gg, WW^*) with less background for detection and we can run at CME equal to m_H for direct production.

Current detector design for FCC-ee is IDEA (Innovative Detector for Electron-positron Accelerator) [2]

- Ultra-light drift chamber for precise tracking and particle identification
- Dual-readout calorimeter for excellent jet energy resolution
- Essential for distinguishing gg from $q\bar{q}$ jets

IV. Machine Learning Approach

Machine learning can **build relationships between the data** within its black box to identifiy which type of event has produced these jets. We can change our model architecture and what we feed the model to see if this improves its detection efficiency.

Two main approaches to ML detection this project will use:

A) Supervised Approach: Representation Learning

Jet Input $\to$ Encoder $\to$ Latent Space $\to$ Classifier $\to$ gg or qq

- Autoencoder learns **compressed jet representation**
- Latent space fed to classifier
- Output: gg vs $q\bar{q}$ classification score

B) Unsupervised Approach: Anomaly Detection

Jet Input $\to$ Encoder $\to$ Latent Space $\to$ Decoder $\to$ Reconstruction

- Learns to encode $q\bar{q}$ jets into latent space then reconstruct them accurately from this space (low loss)
- When feed gg jets should have a bad reconstruction (high loss)
- Doesn't know what gg or $q\bar{q}$ jets are

Performance metrics:
Can use different metrics to measure how well the model is identifying the two jets:

- Gluon tag efficiency
- Quark mistag rate
- ROC Curves

Particle Transformer / ParticleNet:
A new approach for jet tagging where jets are represented as unordered set of its constituent particles (Particle Clouds). A custom neural network architecture called ParticleNet has been developed using attention mechanisms to learn multi-particle correlations. This will be expored for this project. [4, 5]

V. Expected Rewards

Currently, the Yukawa coupling has only been measured for third generation fermions and by the end of the LHCs lifetime only a few second gen (μ and c-quark) will have been studied [1]. This means that, due their small coupling to Higgs field, the mechanism where the stable matter of the visible universe gets their mass from will remain untested and unconfirmed. From this project we hope to assess the feasibility of detecting electron–Yukawa coupling events and develop transferable ML models for future FCC-ee analyses.

References:
[1] Measuring the electron Yukawa coupling, arXiv:2107.02686 [2] Fermilab IDEA Detector, arXiv:2502.21223 [3] fcc-physics-events.web.cern.ch
[4] ParticleNet, arXiv:1902.08570 [5] Particle Transformer ,arXiv:2202.03772