



IMPERIAL



## Multimodal training of jet taggers on subMIT

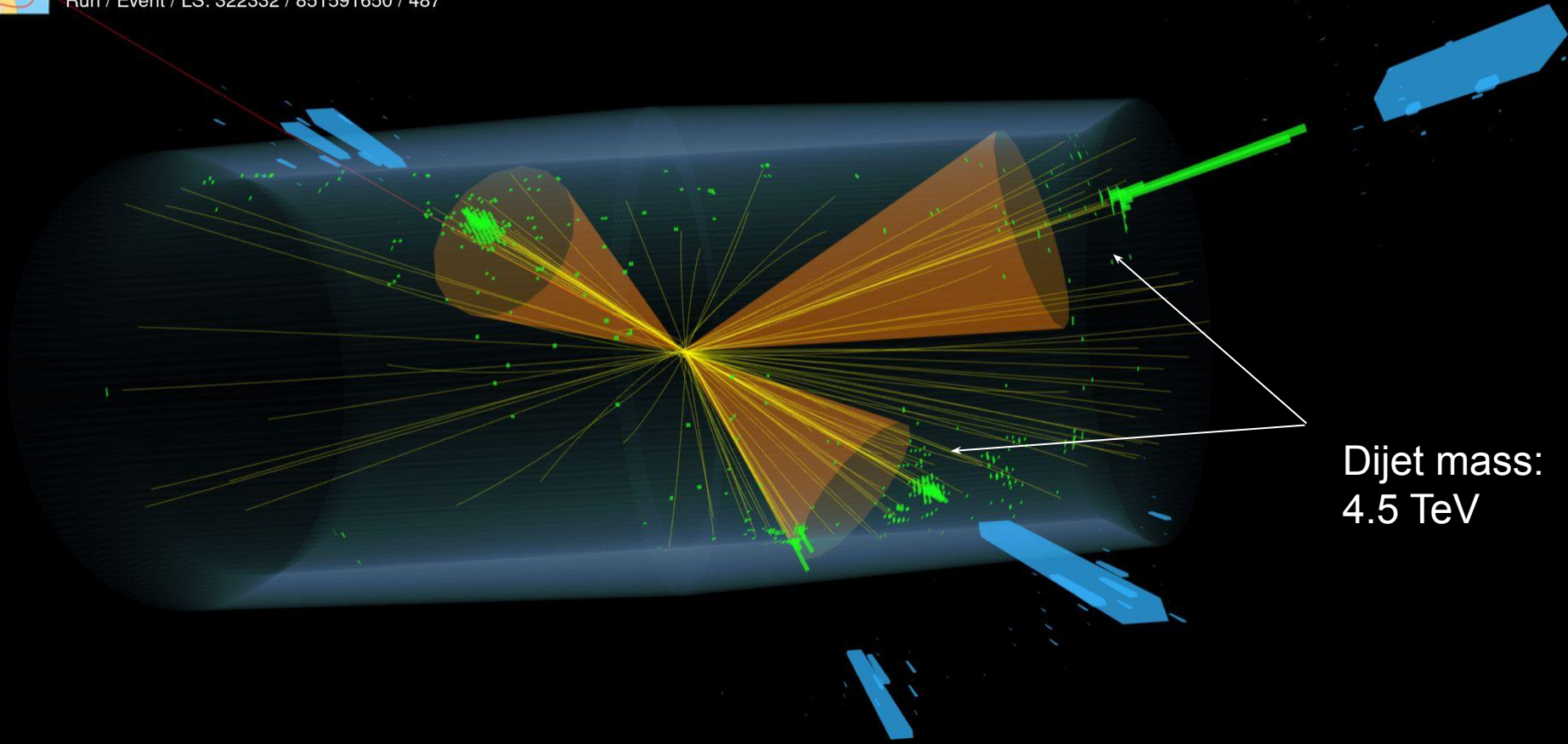
October 27, 2025 | subMIT meeting  
Benedikt Maier (Imperial College London)



CMS Experiment at the LHC, CERN

Data recorded: 2018-Sep-06 05:06:55.343296 GMT

Run / Event / LS: 322332 / 851591650 / 487



Dijet mass:  
4.5 TeV



No Dark Matter candidate in SM

Inferred from **astrophysical observations** like gravitational lensing or rotational curves:

→ Dark matter, **5x more abundant** than visible matter

|         |           |            |          |     |
|---------|-----------|------------|----------|-----|
| $u$     | $c$       | $t$        | $g$      |     |
| $d$     | $s$       | $b$        | $W$      |     |
| $e$     | $\mu$     | $\tau$     | $Z$      |     |
| $\nu_e$ | $\nu_\mu$ | $\nu_\tau$ | $\gamma$ | $H$ |

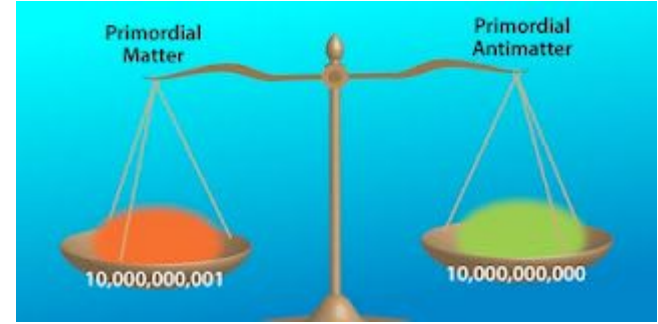
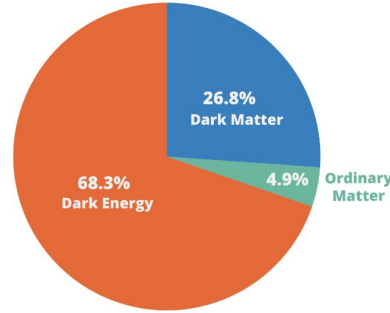


# Many open questions in cosmology and particle physics

## Experiment-driven:

Dark Matter & Dark Energy

Matter-antimatter asymmetry



## Theory-driven:

Hierarchy problems (weakness of gravity, fine tuning at level  $10^{16}$ )

Why 3 generations of fermions?

Dark sector could hold **the answer** to many of these questions



quantumdiaries.org

→ Probing it with jets

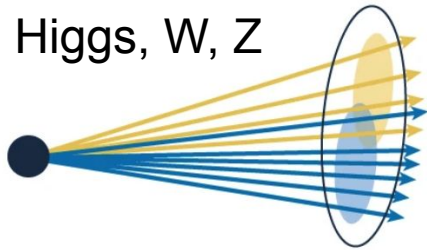
# Characterizing jets

abundant at LHC

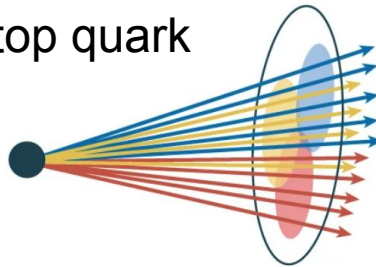
quark, gluon



Higgs, W, Z



top quark

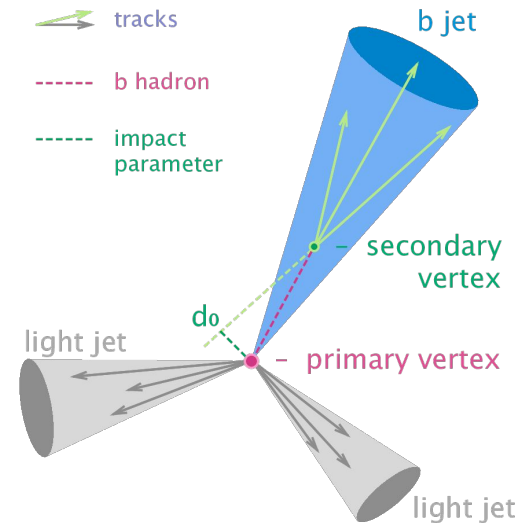


Different particles → different substructure

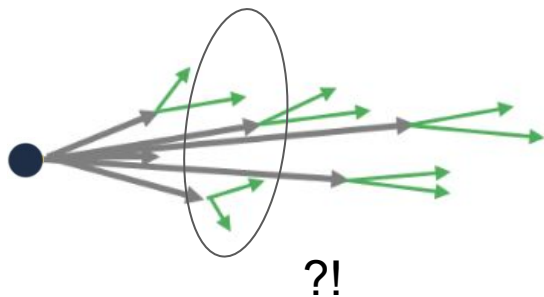
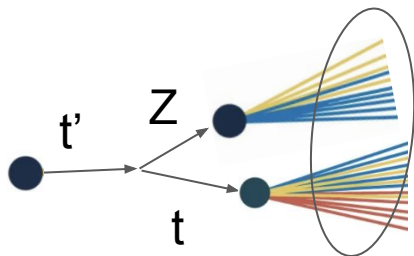
→ Exploit substructure through algorithms targeting e.g., "pronginess"

B hadrons long-lived  
→ displaced "secondary" vertex within jet

→ Exploit flavor content  
(e.g., multivariate B tagging algorithms)



# Anomalous jets



- **New** physics could have:
  - long decay chains of SM resonances
  - completely exotic particles & weird radiation patterns
- Jet tagging potentially a key to new physics

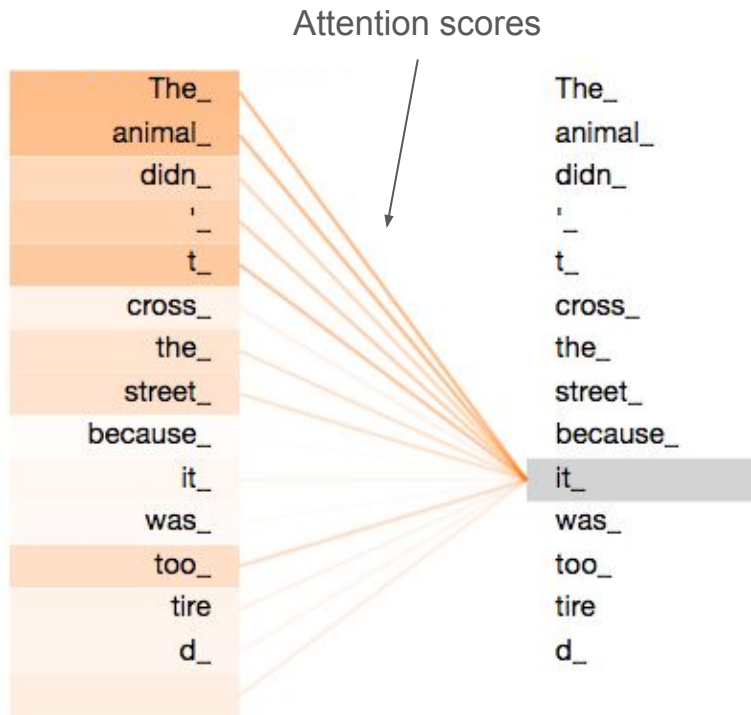
# Attention!

Use **transformers** (same as ChatGPT) to *translate* sequence of particles into sequence of pileup probabilities

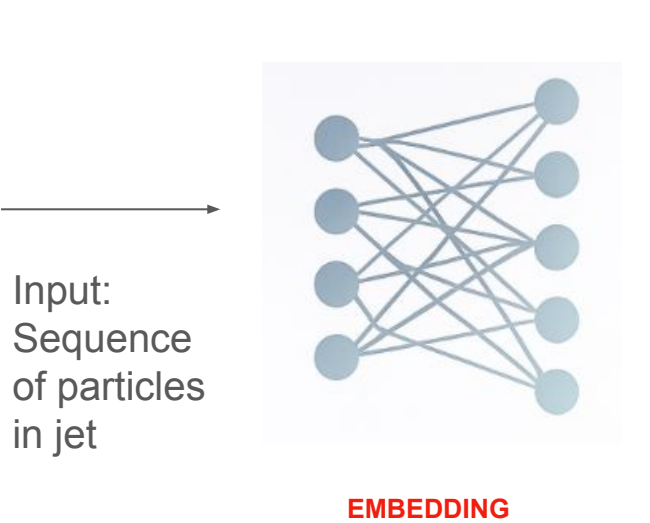
Key concept called **Attention**

→ Particles “talk” to each other to see how much they have in common

→ Very efficient extraction of relevant physics information



# How every state-of-the-art tagger looks like



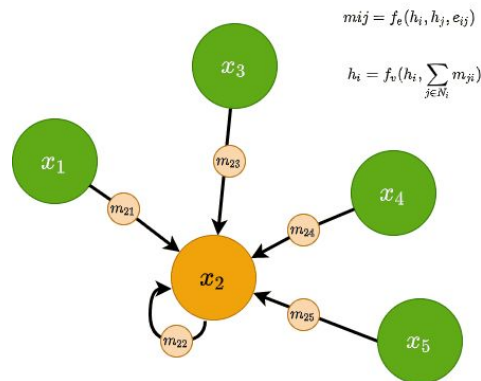
Projecting particles into a  
higher-dimensional  
(abstract, "latent") space

# How every state-of-the-art tagger looks like

Input:  
Sequence  
of particles  
in jet



**EMBEDDING**



**MESSAGE PASSING**

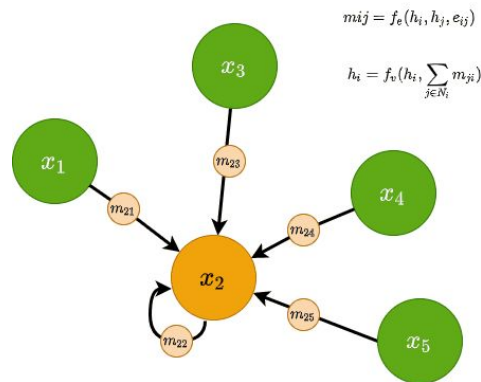
2D convolution  
Graph convolution  
Attention, ...

# How every state-of-the-art tagger looks like

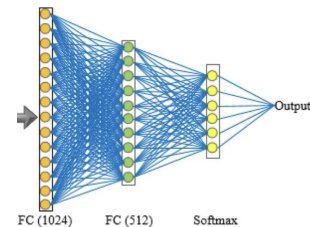
Input:  
Sequence of particles  
in jet



**EMBEDDING**



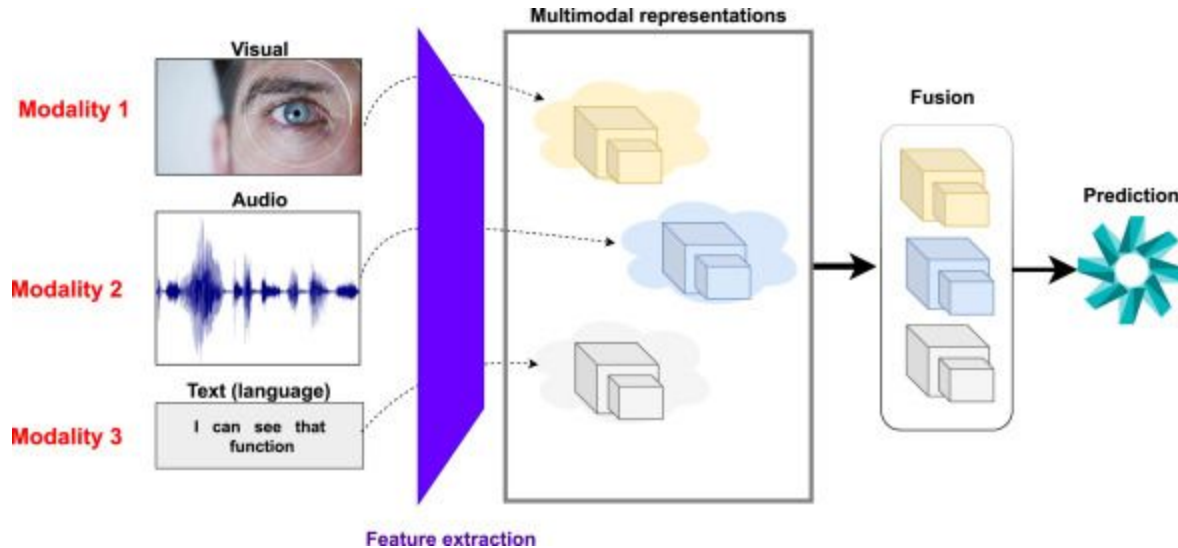
**MESSAGE PASSING**



**GLOBAL POOLING/SOFTMAX**

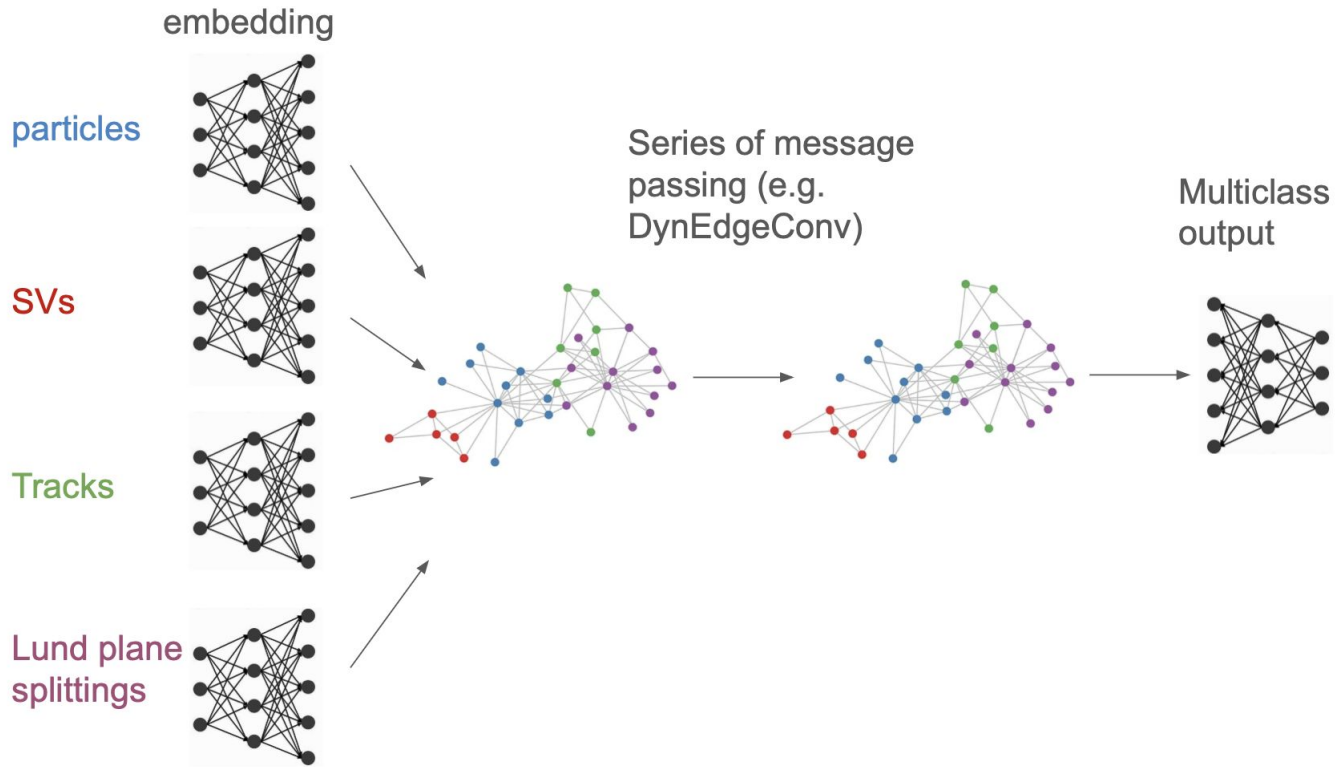
Aggregating information  
to go from particle-level  
to jet-level (e.g., final  
class label for jet)

# Multimodal inputs

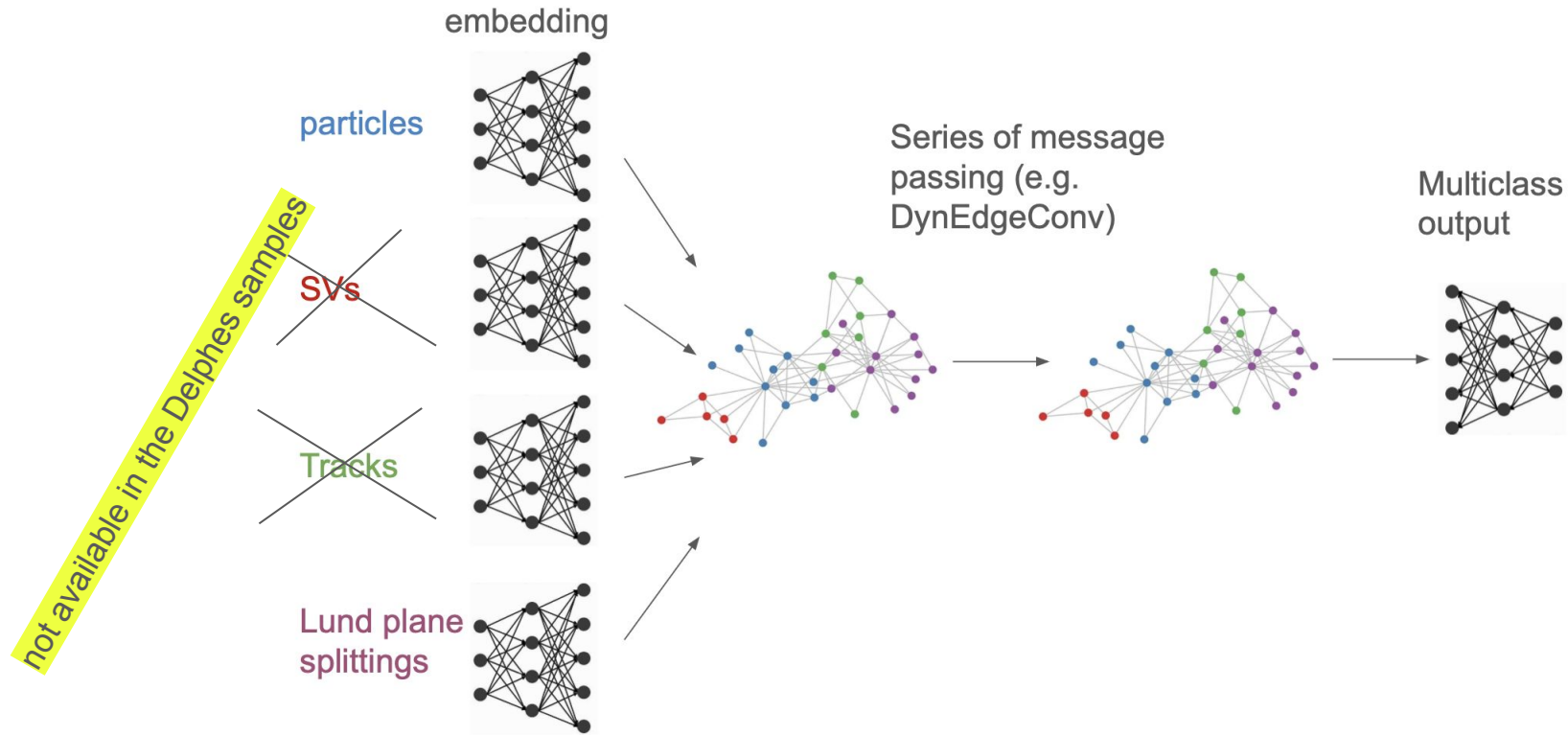


Helps the network to identify context and relations within the network

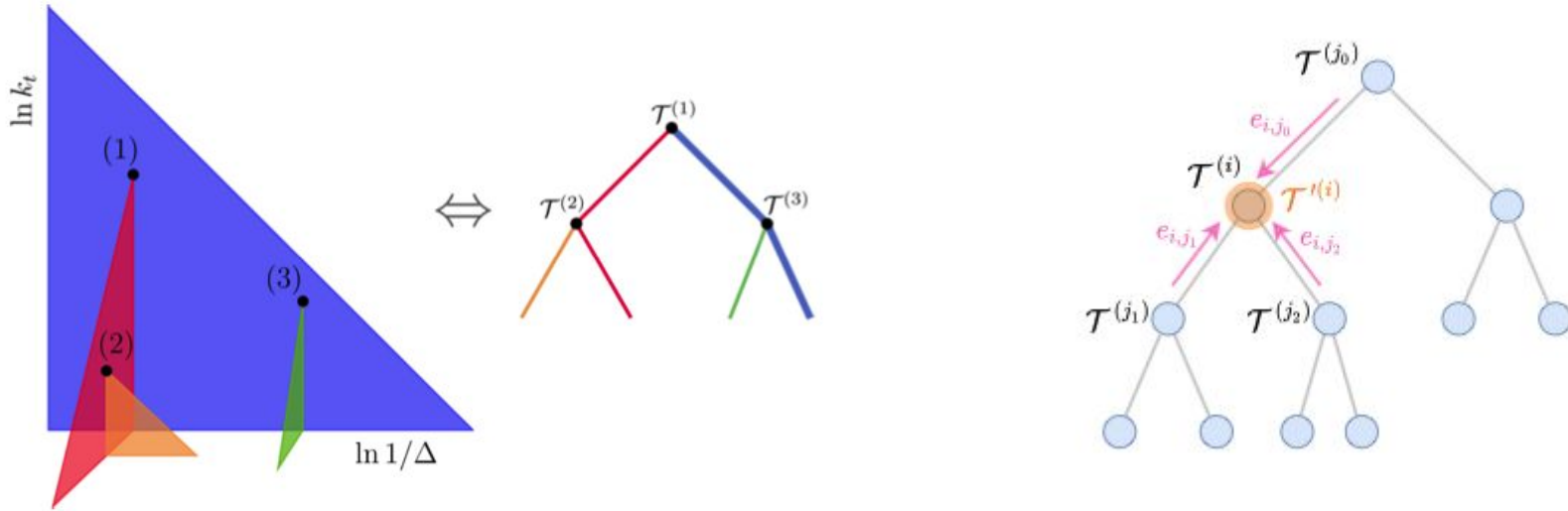
# Multimodal inputs



# Multimodal inputs

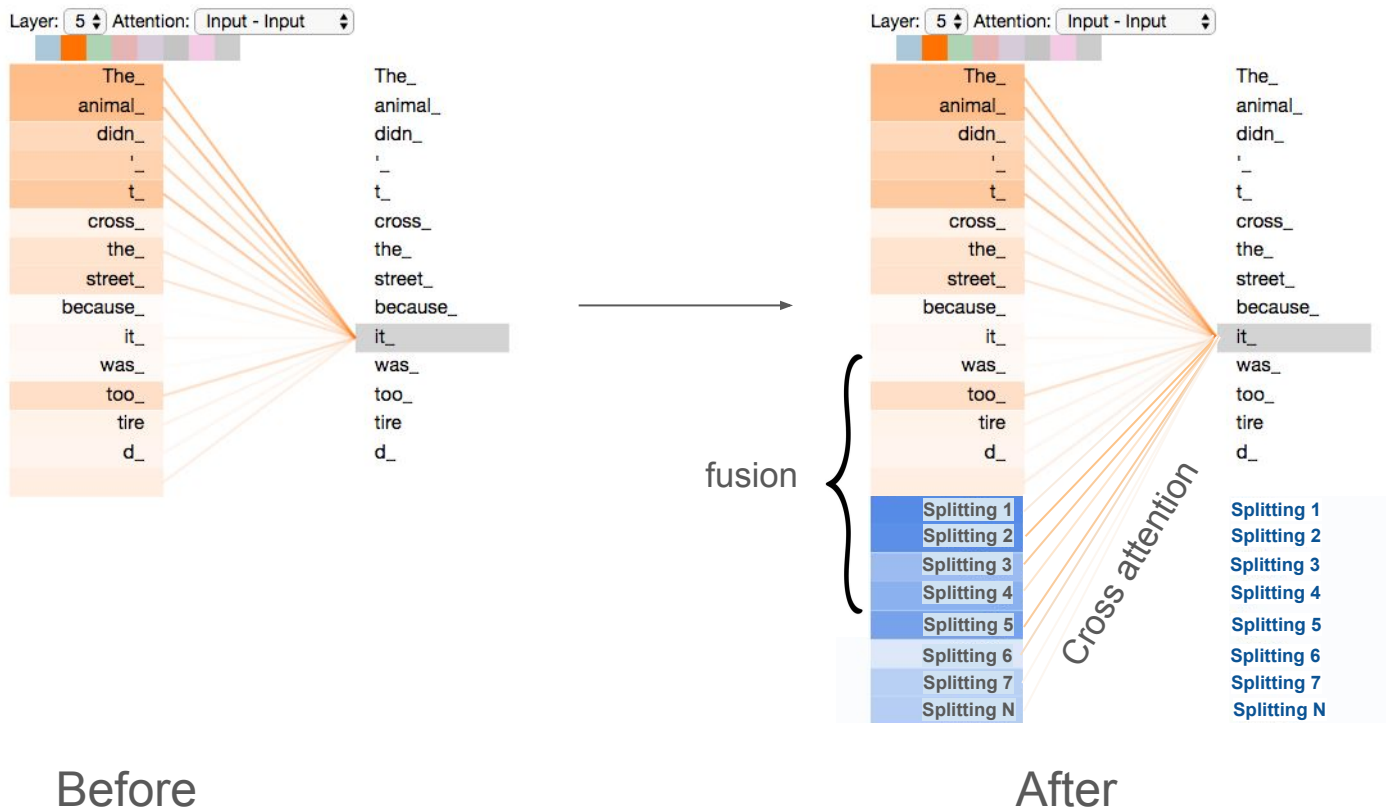


# The Lund jet plane

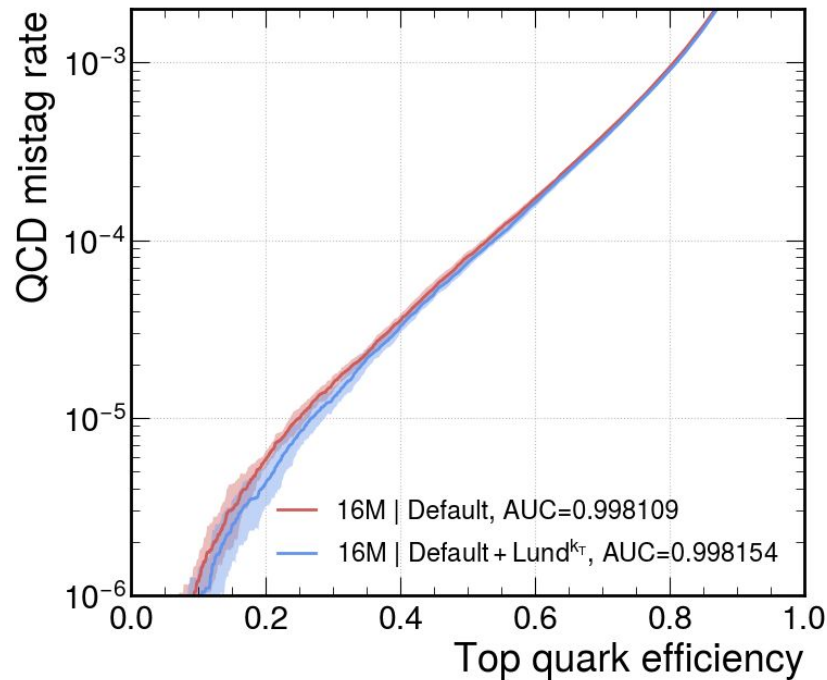
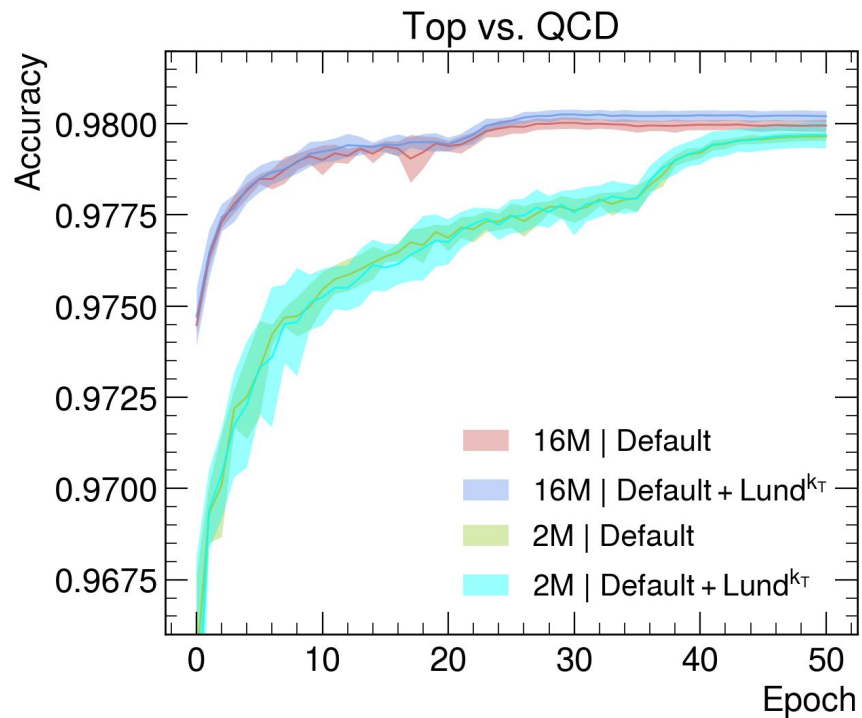


- **Lund Plane:** A two-dimensional phase space that visualizes the branching structure of a jet, plotted in terms of the splitting angle ( $\Delta R$ ) and relative transverse momentum ( $k_\perp$ ) of emitted particles.
- **Purpose:** Provides an interpretable map of the jet's internal structure, linking measured emissions to QCD radiation patterns and enabling comparisons between data and theory.
- **Lund Plane Splittings:** Each emission (or "splitting") in the jet corresponds to a point on the plane, allowing the entire cascade of parton branchings to be represented as a graph that encodes the jet's substructure evolution.

# Adding the Lund plane splittings: Cross-attention fusion



# Results

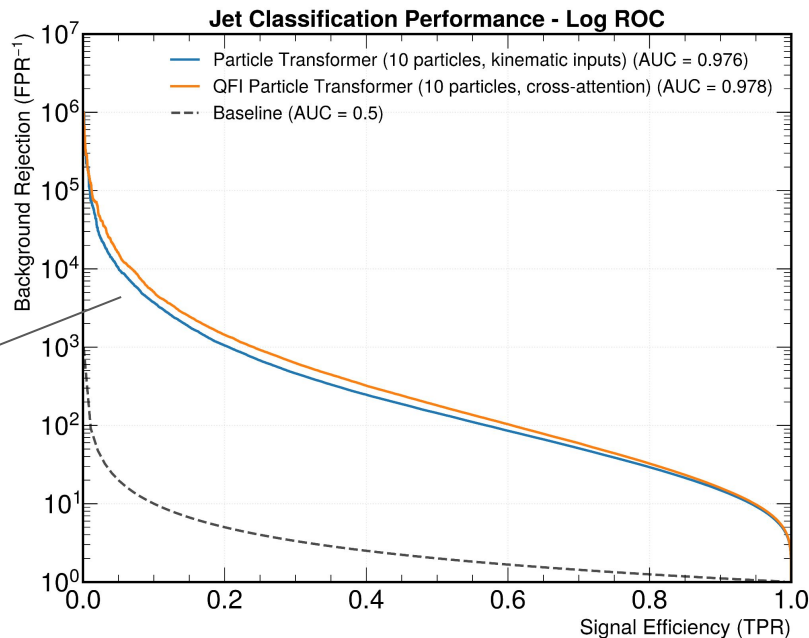


# More results to come

How does it look for other heavy resonance decays:  $H \rightarrow cc$  vs QCD,  $H \rightarrow 4q$  vs QCD ...

What other modalities can one use?  
Currently exploring Quantum Fisher information (a quantum view of the jet)

...



# What I found out working on subMIT

- Environment installed in my /work/submit dir, activated from within job
- submit20,21,22... much faster than all the others
- One epoch (20M jets) takes roughly half a day vs. a full day on submit70...
- Streaming data (several hundred GBs) from /scratch/ much faster than from /ceph/
- Have not explored yet multi-GPU training