



UNIVERSITÄT
HEIDELBERG
ZUKUNFT
SEIT 1386

EuCAIFCon 2024

Weak supervision for quark/gluon tagging in CMS Open Data

(Poster #81)

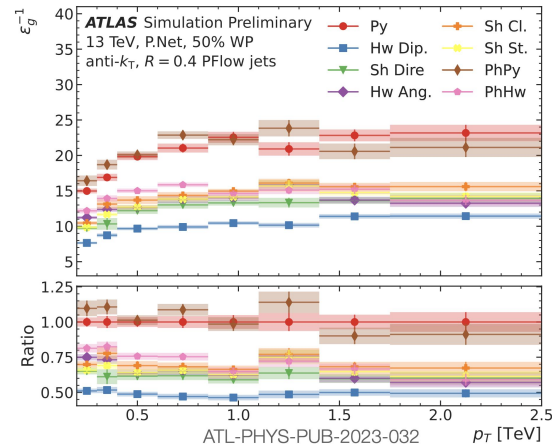
Ayodele Ore

In collaboration with Matthew J. Dolan and John Gargalionis

Weakly-supervised quark/gluon tagging

- At LHC, high-energy quarks and gluons both lead to *jets*.
- Neural networks are good classifiers, but full supervision inherits large theory uncertainty from simulation.
- Weak supervision allows models to be trained directly on data, using **mixed quark/gluon** samples.

Metodiev et al.
JHEP10(2017)174



Full supervision

Weak supervision

(Quark jets) **QQQQQ**

QQQGG (Z + jet events)

VS

VS

(Gluon jets) **GGGGG**

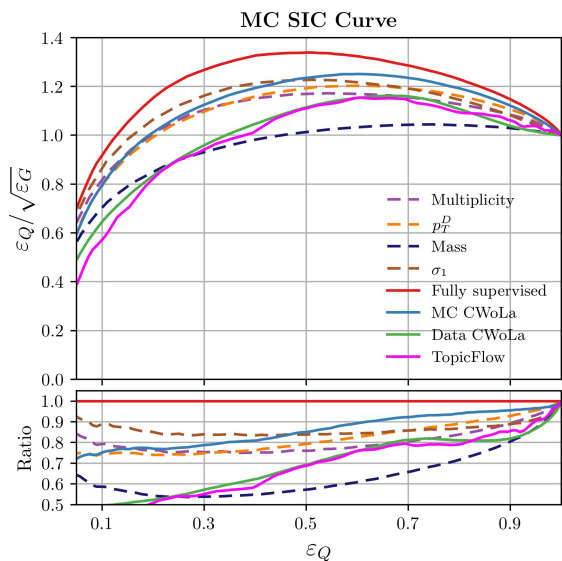
QGGGG (Dijet events)

Can isolate in data

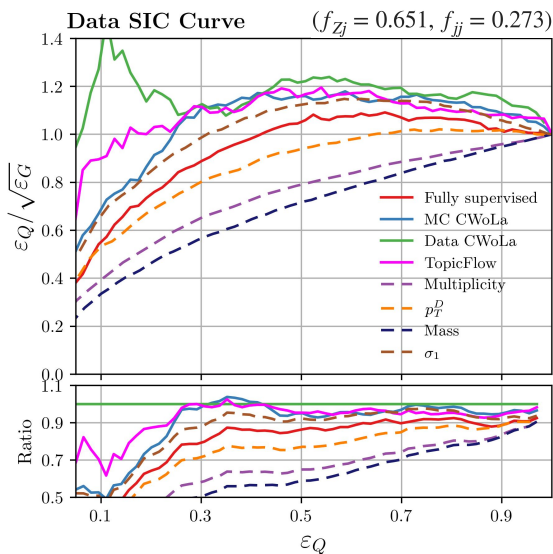
Our study: How do models rank on real data?

1. Estimate mixture fractions

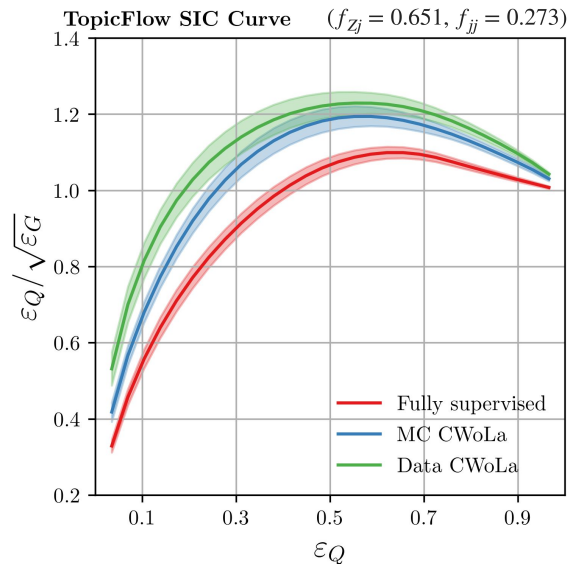
2. Train TopicFlow



Simulation



Data



Generative model