CHEP 2026
Advancing the CMS Level-1 Trigger: Jet Tagging and p_T regression with Deep Sets at the HL-LHC

Stella Schaefer, Oz Amram, Christopher Brown, Duc Minh Hoang, Sioni Summers, Sebastian Wuchterl
for the CMS Collaboration

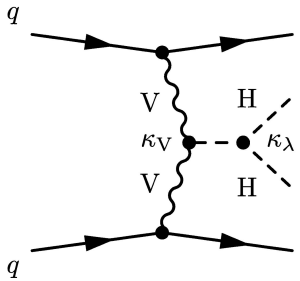
CERN | NexTGen Next Generation Triggers

L1 Trigger - rate reduction through event rejection

Phase-2 (High-Lumi LHC):

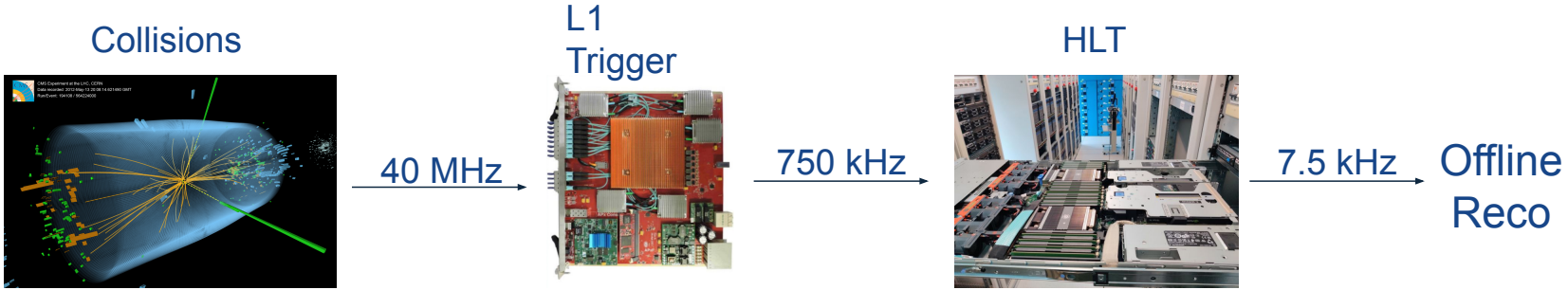
- **Currently: 60 simultaneous interactions (PU)**
  - **→ Phase-2: up to 200 PU**
- **First time: (outer) tracker in L1 (p_T > 2 GeV)**

➔ **New seeds / algorithms possible**



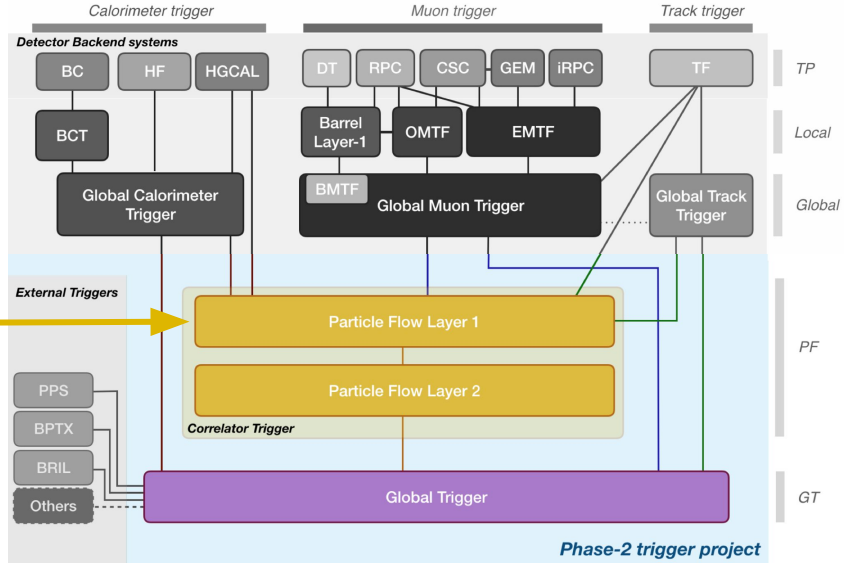
Requirements:

- **Fast (12.5 μs)**
- **Run on FPGAs**



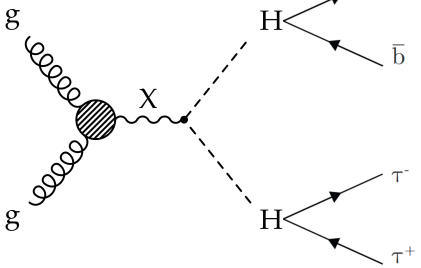
Jet Tagging at L1

- Reconstruct up to 16 jets per event at L1
- Jet tagging to identify flavour
- Helps identify important physics processes, e.g:
  - Di-Higgs
  - Vector boson fusion (VBF)
- Implementation in correlator trigger

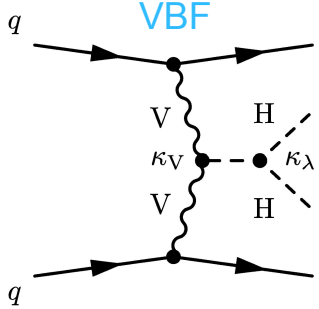


Phase-2 Report

Di-Higgs



VBF



Jet Energy Corrections at L1

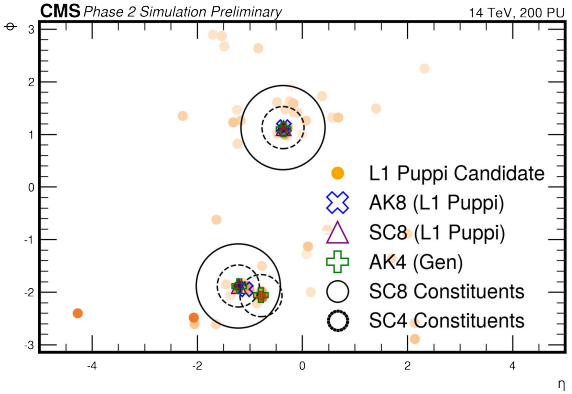
- Jets from Seeded Cone (SC) algorithm
  - SC performance similar to anti-kt
- JECs through fit of the response (η, p_T binned)
- Fitting function per η bin

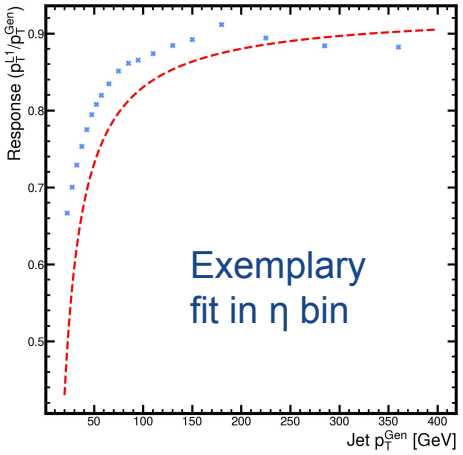
$$R = \frac{p_{T,reco}}{p_{T,gen}} = a + \frac{b}{p_{T,gen}}$$

$$p_{T,reco} = a \cdot p_{T,gen} + b$$

$$p_{T,corr} \equiv p_{T,gen} \approx \frac{1}{a}(p_{T,reco} - b)$$

➔ Linear function to derive $p_{T,corr}$

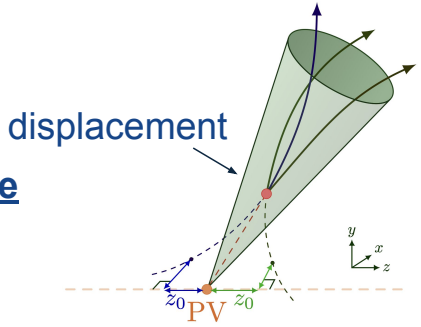


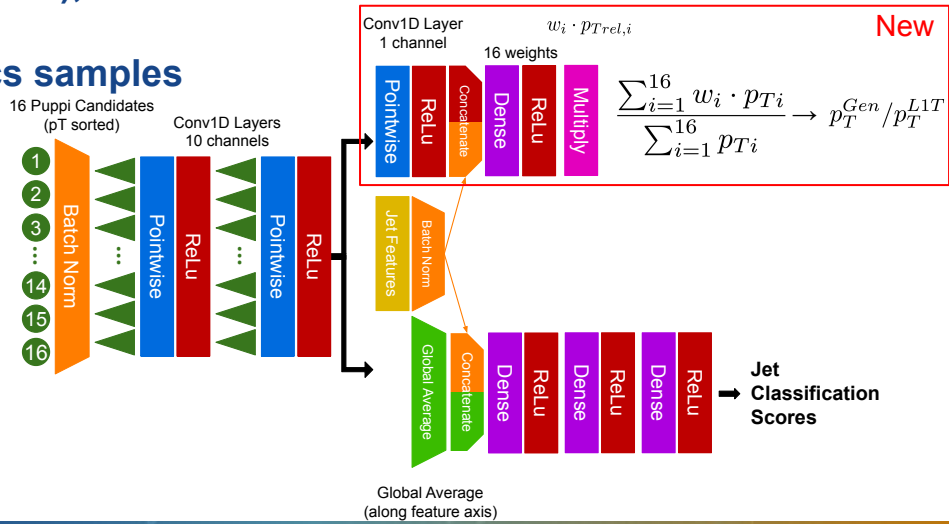


New Strategy for Phase-2

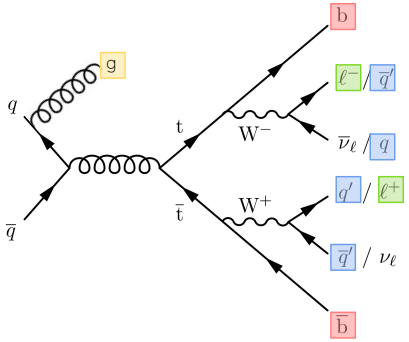
- Multiclass model for
  - Flavour tagging (b, c, light, gluon, τ +/- , e, μ) + new pile-up score
  - p_T regression (flavour aware)
- 16 leading p_T constituents (of SC Jets)
- Inputs:
  - Constituents: kinematic, particle IDs (PIDs), “record filled”
  - Jet: p_T, η
- Trained on a mixture of Monte Carlo physics samples
  - kinematic cuts p_T > 15 GeV, |η|<2.4
  - Displaced tracking helps classification
- Small, quantized model necessary
- Hardware optimised architecture

➔ Deep Sets



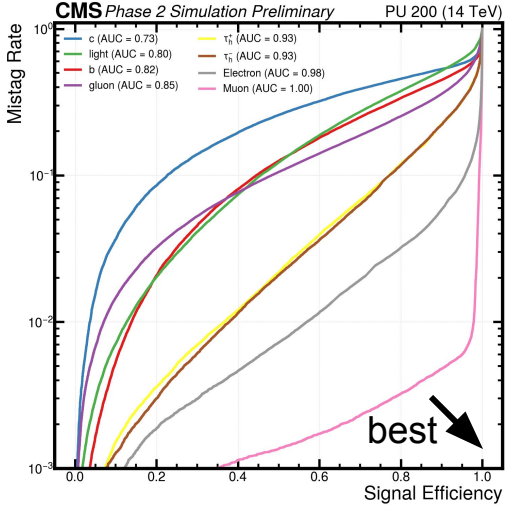


Flavour Tagging

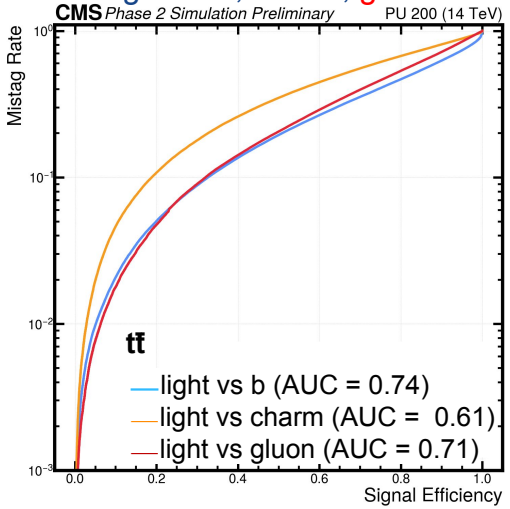


- Important one-vs-one classifiers
  - Light vs gluon for VBF
  - b vs light for Higgs
- $t\bar{t}$ can contain leptons, quarks and gluons in final state

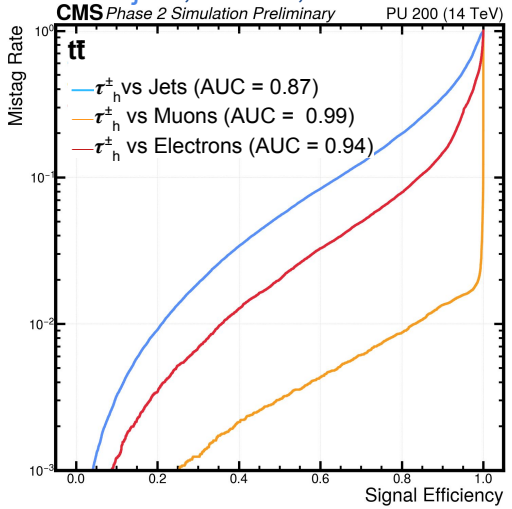
one vs. all



light vs b, charm, gluon



τ vs jets, muons, electrons

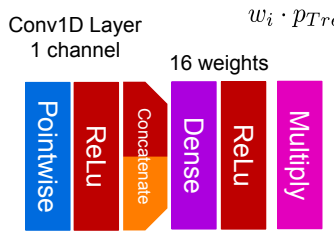


gluon
b
light
Electron
Muon
τ_h^-
τ_h^+
c

➔ Full tagging developments here

p_T Regression

- Derivation of multiplicative correction factor
- Jet correction through constituent corrections
  - p_T weight per constituent
- Weighted average to derive jet correction



$w_i \cdot p_{Trel,i}$

$$\frac{\sum_{i=1}^{16} w_i \cdot p_{Ti}}{\sum_{i=1}^{16} p_{Ti}} \rightarrow p_T^{Gen}/p_T^{L1T}$$

Sample weights

- Flat p_T distribution per class multiplied
- p_T weighting

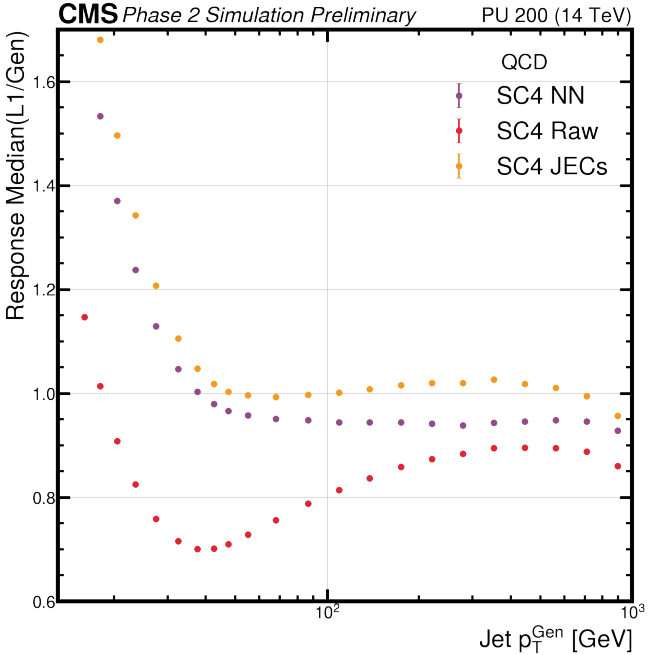
Loss weights for Huber Loss per jet (HL_{jet})

- Punish underestimation
- Punish overestimated pile-up jets
- ➔ Tunability through α, pu factors

$$(\alpha \cdot |y_{true} - y_{pred}| + 1) \cdot HL_{jet}$$
$$(pu \cdot |1 - y_{pred}| + 1) \cdot HL_{jet}$$

Response

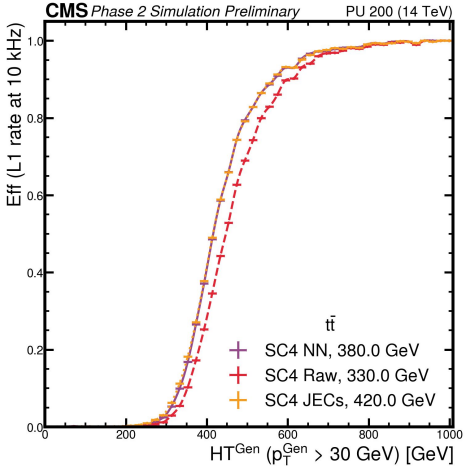
- Response for median per p_T bin
- JECs:
  - Optimization of median
  - Approach 1 from above
- NN jets:
  - Optimization of mean
  - Approach 1 from below

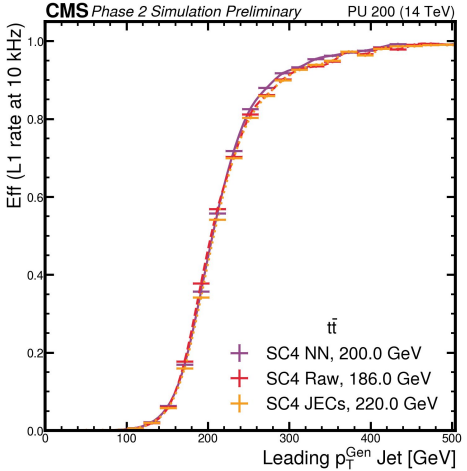


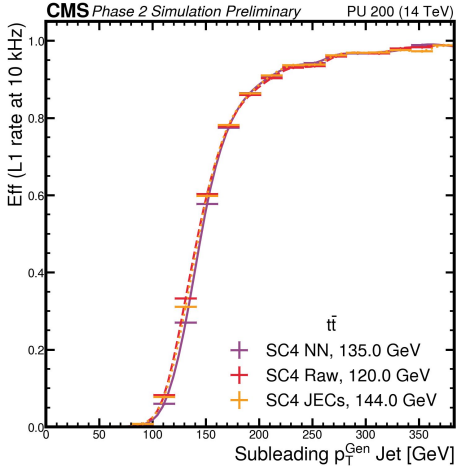
Turn-on Curves

- Derived for $t\bar{t}$
- Turn-on at fixed rate of 10 kHz
  - Rate estimation on minimum bias
- Derived for
  - HT (=∑ jet p_T)
  - Leading jet p_T
  - Subleading jet p_T (ie. symmetric dijet trigger)

SC4 NN
SC4 Raw
SC4 Jecs





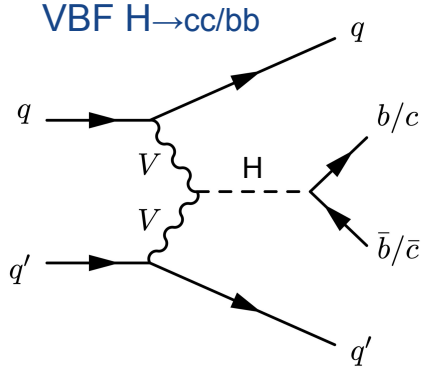


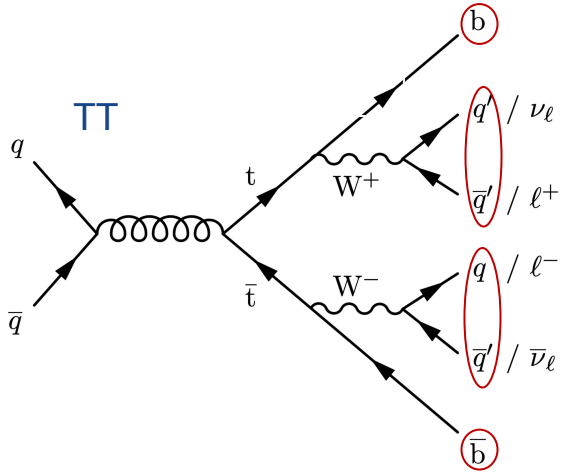
Invariant Mass Reconstruction

Reconstructed invariant masses of

- **Hadronically decaying top quark and W boson**
  - **Derived on $t\bar{t}$**
- **Higgs boson decaying to two**
  - **Charm quarks (VBF H→cc)**
  - **Bottom quarks (VBF H→bb)**

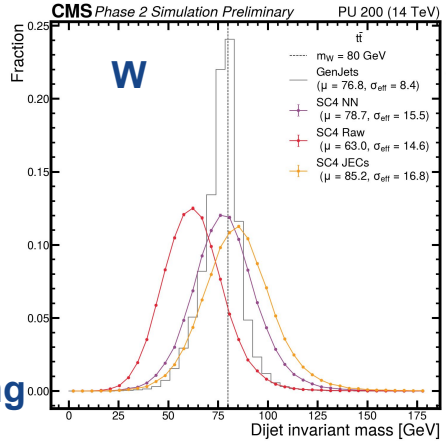
- **Decay products identified through ΔR matching**
- **Peak width σ_{eff}: smallest 68% interval**

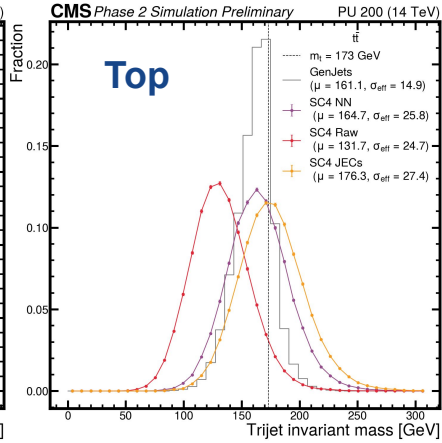


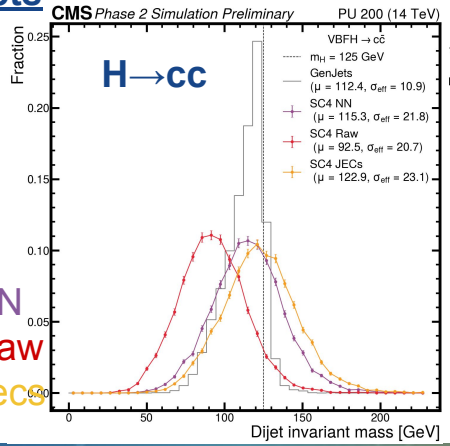


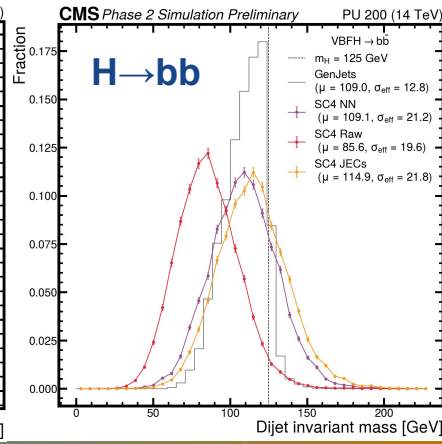
Invariant Mass Reconstruction

- Hadronically decaying top quark and W boson
  - Derived on $t\bar{t}$
- Higgs boson decaying to two
  - Charm quarks (VBF H→cc)
  - Bottom quarks (VBF H→bb)
- Decay products identifies through ΔR matching
- Peak width σ_{eff}: smallest 68% interval
- Improved peak position relative to generator jets
- Reduced peak width





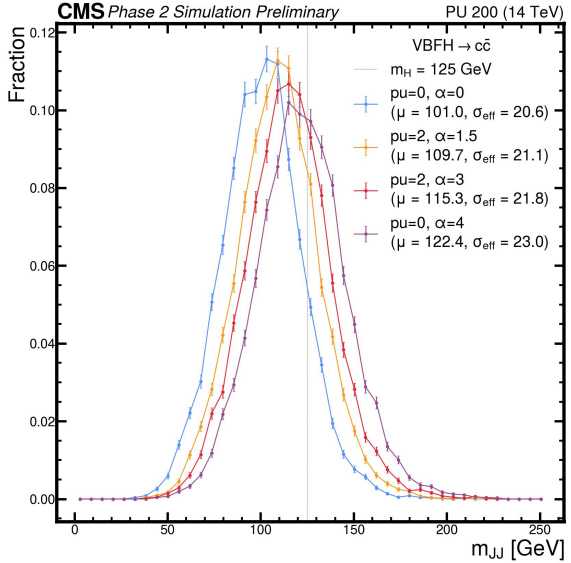




SC4 NN
SC4 Raw
SC4 Jecs

Mass Peak Calibration

- **"Default model", ie. only p_T scaling of samples**
  - **Undercorrection of p_T**
- **Punish underestimation in the loss to calibrate peak position** $(\alpha \cdot |y_{true} - y_{pred}| + 1) \cdot HL_{jet}$
  - **At the cost of increasing peak width**



Summary

- Deep Sets Multiclass Tagger
- Compromise of performance and latency / resources
- 9 tagging classes
- p_T regression
  - Individual constituent p_T weights
- Improved and tunable p_T calibration
  - Reconstructed invariant masses
- Similar performance for fixed rate turn-ons

Next steps:

- Replace JECs with NN
- Leverage new model features in global trigger

