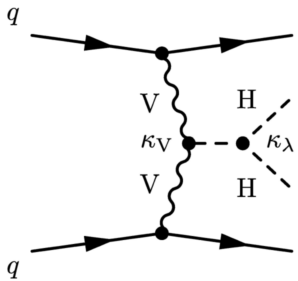
ML4Jets, 17.09.2026
Advancing the CMS Level-1 Trigger: Jet Tagging with DeepSets at the HL-LHC

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

L1 Trigger - rate reduction through event rejection



Phase-2 (High-Lumi LHC):

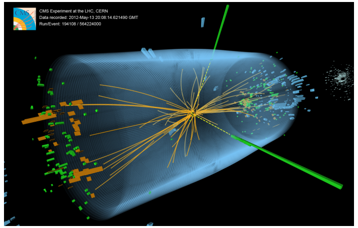
- **Currently: 60 simultaneous interactions (PU)**
  - **→ [Phase-2](): up to 200 PU**
- **First time: (outer) tracker in L1 (p_T > 2 GeV)**

➔ **New seeds / algorithms possible**

Requirements:

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

Collisions



40 MHz

L1 Trigger

750 kHz

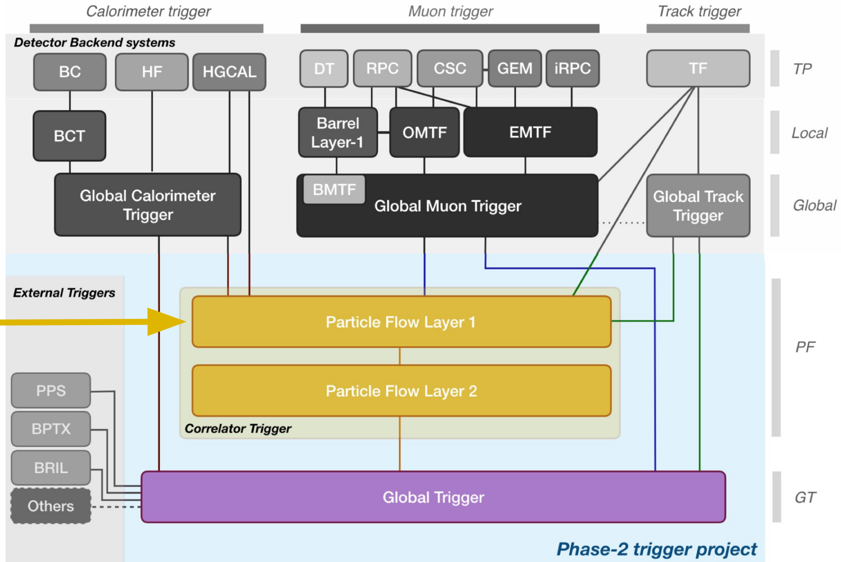
HLT

7.5 kHz

Offline Reco

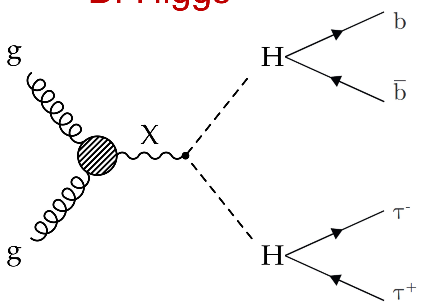
Jet Tagging at L1

- 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

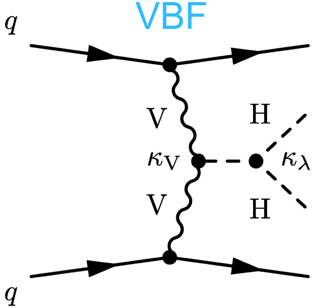


Phase-2 Report (CMS-TDR-021)

Di-Higgs

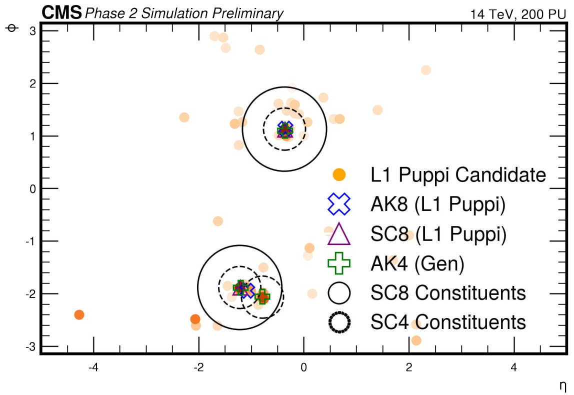


VBF



Jet Energy Corrections at L1

Proceedings CHEP 2023



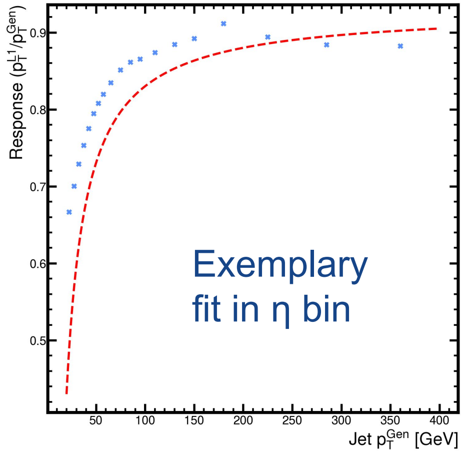
- Jets from Seeded Cone (SC) algorithm
  - SC performance similar to anti-kt
- JECs through fit of the response (η, p_T binned)
- Fitting function per η bin

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

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

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

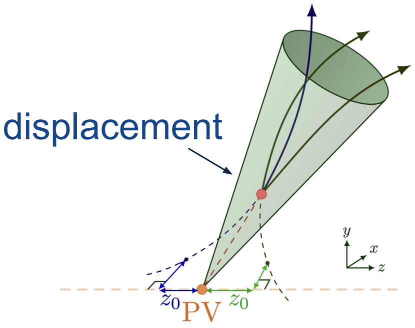
➔ Linear function to derive $p_{T,corr}$



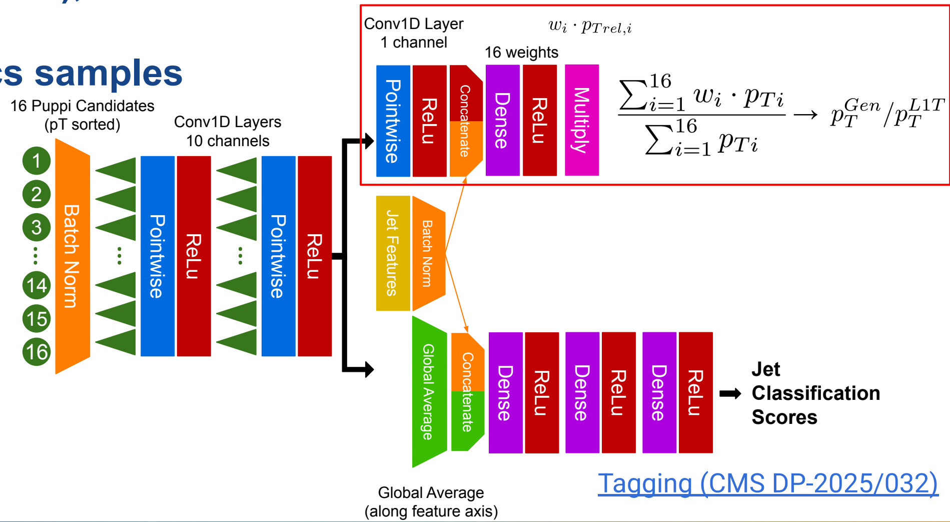
New Strategy for Phase-2

- Multiclass model for
  - Flavour tagging (b, c, light, gluon, τ +/- , e, μ) + new PU 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



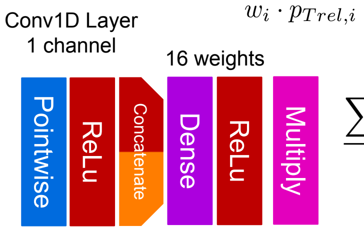
Regression (CMS DP-2026/037)



Tagging (CMS DP-2025/032)

p_T Regression

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



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

Sample weights

- Flat p_T distribution per class multiplied
- p_T weighting

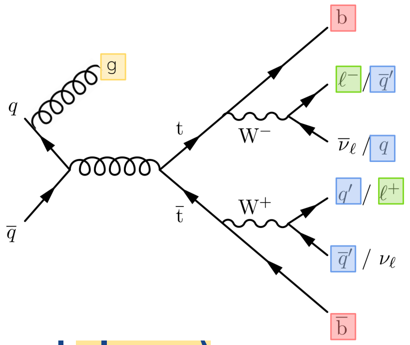
Loss weights for Huber Loss per jet (HL_{jet})

- Punish underestimation
- Punish overestimated pile-up jets

➔ Tunability through α, pu factors

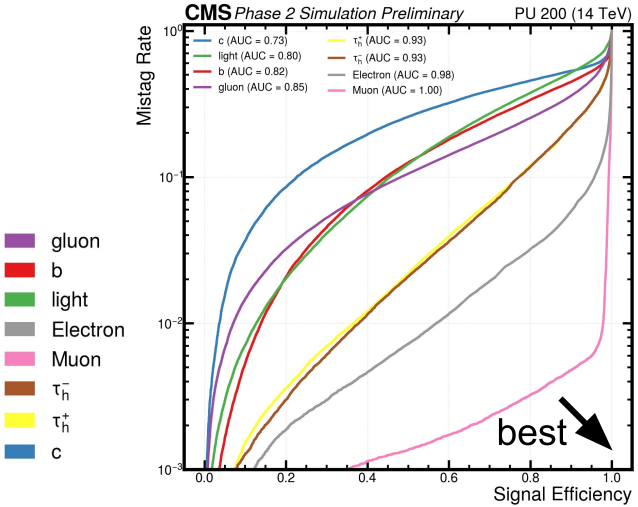
$$(\alpha \cdot |y_{true} - y_{pred}| + 1) \cdot HL_{jet}$$
$$(pu \cdot |1 - y_{pred}| + 1) \cdot HL_{jet}$$

Flavour Tagging

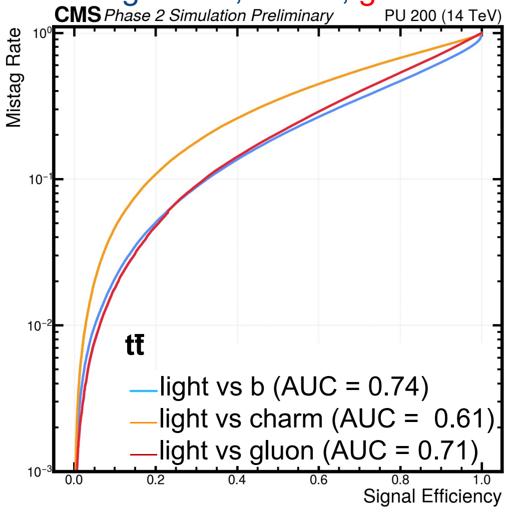


- **Important one-vs-one classifiers**
  - **Light vs gluon for VBF**
  - **b vs light for Higgs**
- **$t\bar{t}$ can contain all tagged flavours in final state (leptons, quarks and gluons)**

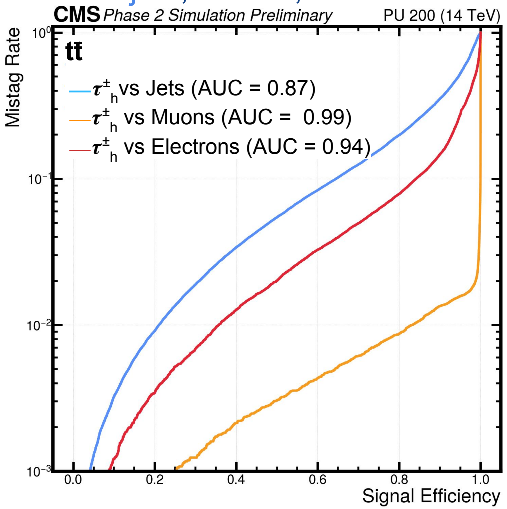
one vs. all



light vs b, charm, gluon



τ vs jets, muons, electrons



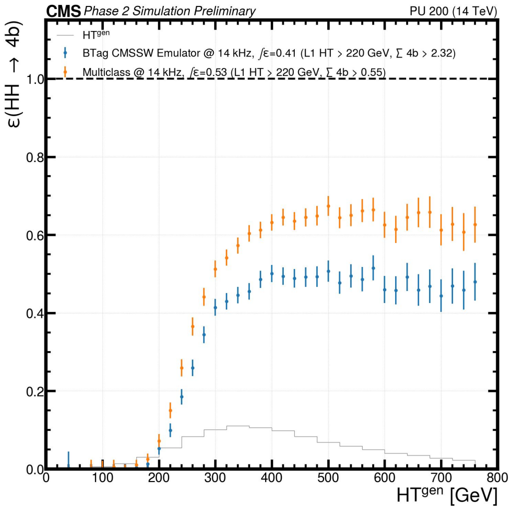
➔ **Full tagging developments [here (CMS DP-2025/032)]**

$H \rightarrow$bbbb / bb$\tau\tau$

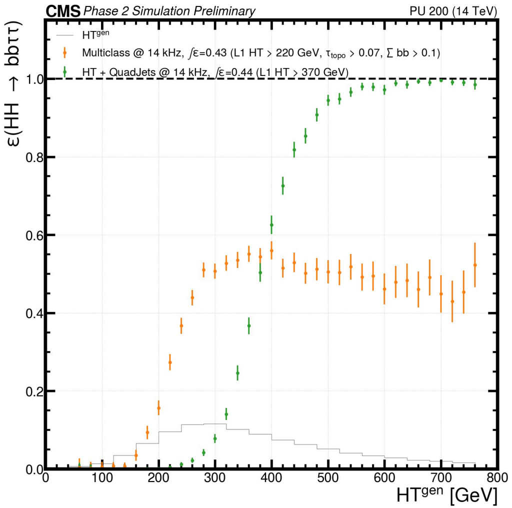
Identification of HH → 4b/bb$\tau\tau$

- Binary b tagging + H_T (220 GeV)
- Multiclass jet tagging
- Comparison at same rate ie.:
  - number of selected events
- Overall gain in efficiency
  - important gain in low H_T
- Before: No tagging (H_T threshold)
- Multiclass jet tagging
  - Lower fixed H_T threshold

[Tagging (CMS DP-2025/032)]



$$\tau_{\text{topology}} = \left(p^1_{\tau^+} + p^2_{\tau^+}\right) \times \left(p^1_{\tau^-} + p^2_{\tau^-}\right)$$



p_T regression

Turn-on curves

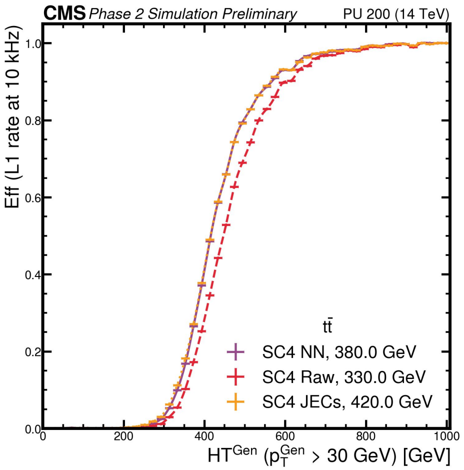
- Derived for $t\bar{t}$
- Turn-on at fixed rate of 10 kHz
  - Rate estimation on minimum bias
- Derived for
  - H_T (=∑ jet p_T)
  - Leading jet p_T

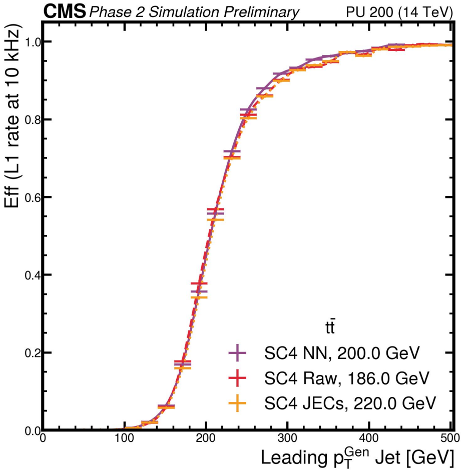
SC4 NN
SC4 Raw
SC4 Jecs

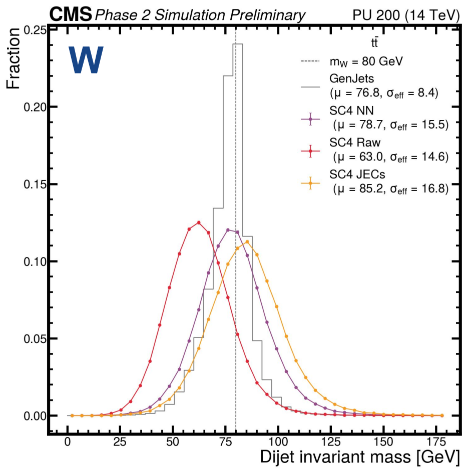
Mass reconstruction

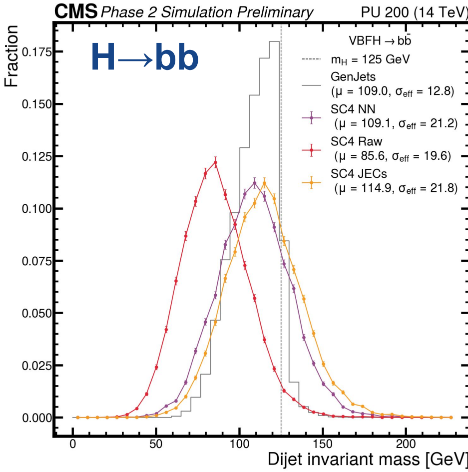
- Reconstructed invariant masses of
  - W boson (hadronically decaying): $t\bar{t}$
  - Higgs: VBF H→bb
- Improved peak position relative to gen jets
- Reduced peak width

CMS DP-2026/037







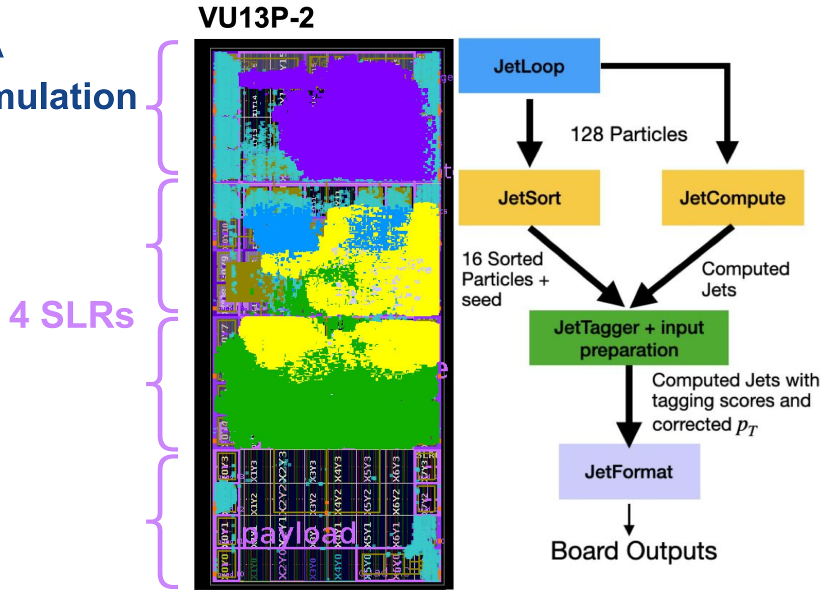


FPGA implementation

- ✅ **Latency and Initiation intervals almost compatible with L1**
  - **155ns (NN) + 33ns (Input preparation) latency taking input at 360 MHz**
- ✅ **NN architecture fits AMD Xilinx VU13P-2 FPGA**
- ✅ **Firmware results match bit-to-bit with CMS simulation software**

	II (clock)	latency	BRAM	DSP	FF	LUT
Jet Tagger	1	155.6 ns	0.11	11.49	5.37	9.74
Input Preparation	1	33.3 ns	0.71	0.13	0.23	0.42

% of VU13P

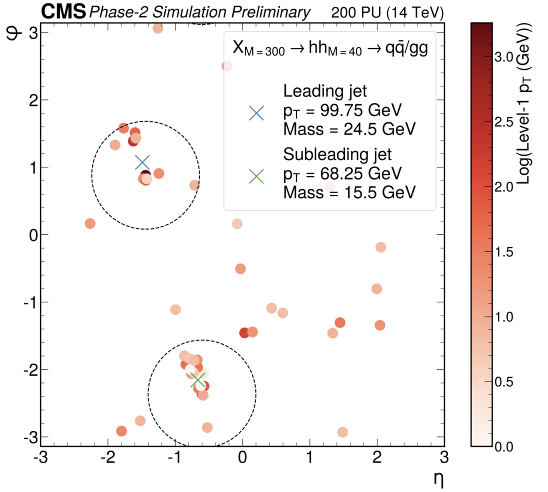


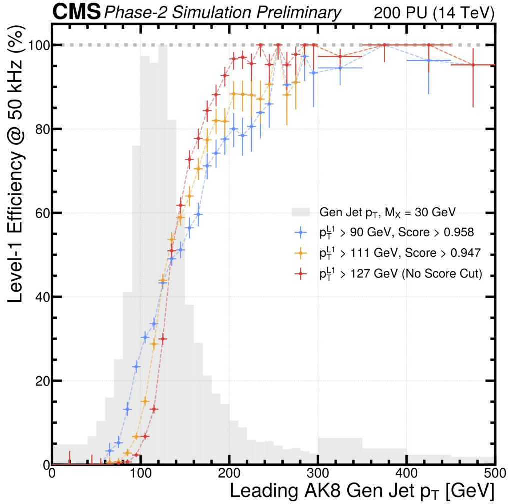
SC8 Jets

- reconstruct collimated decay in a single jet
- mass close to mother particle
- SC8 jet reconstructed in addition to SC4

Tagging CMS-DP-2026-110

- analyze substructure of SC8 jets
  - identify boosted resonance
- similar deep sets architecture
- binary tagging score
- recovered low mass phase space





Summary + Outlook

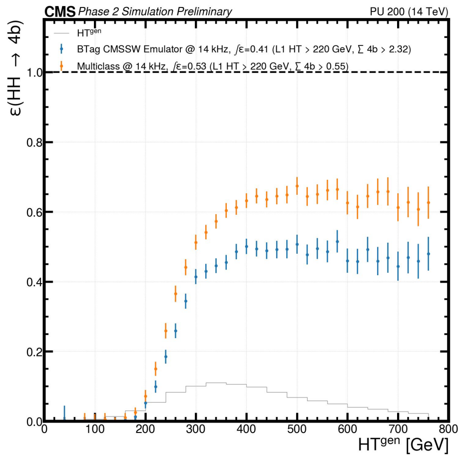
ML for CMS Phase-2

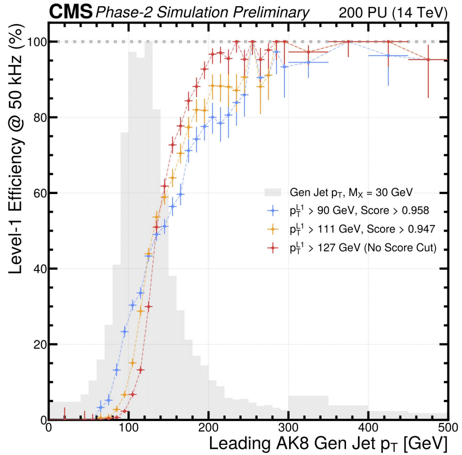
Jet Flavour Tagging

- Triggers for di-Higgs, VBF, …
  - Extensive jet descriptions necessary
- Deep Sets Multiclass Tagger
  - Compromise of performance and latency / resources
- Significant b-tagging improvements
  - HH → 4b/bbττ
- p_T regression improvements
  - invariant mass reconstruction

2-prong Tagging

- collimated decay in SC8 jet
- identify boosted resonance





stella.felice.schaefer@cern.ch