Electron Reconstruction and Identification in the CMS Phase-2 Level-1 Trigger

The CMS Collaboration
cms-dpg-conveners-l1t@cern.ch

Introduction

The CMS Level-1 Trigger (L1T) for the HL-LHC will use powerful FPGA processors connected by a high-bandwidth network of optical fibers. The new system will have access to online tracking, which enables the identification of calorimeter deposits originating from electrons. The Phase-2 L1T system will also access three-dimensional calorimeter information with a high transverse and longitudinal granularity from high-granularity calorimeter (HGCal) covering forward regions. The combination of these new features available at the Phase-2 L1T system will allow for good physics acceptance in harsh HL-LHC environment, identifying physics objects with reduced rates and maintaining trigger thresholds as low as those from the LHC Run 3. This DP note illustrates the current status and performance of the electron identification algorithms, focusing in particular on the new approaches that, exploiting Machine-Learning inference in the FPGAs, enhance the identification efficiency for low transverse momentum electrons, compared to what was presented in the Level-1 Phase-2 TDR [1] and CMS-DP-2023-047 [2].

HGCal Cluster Identification: Multi-class ID
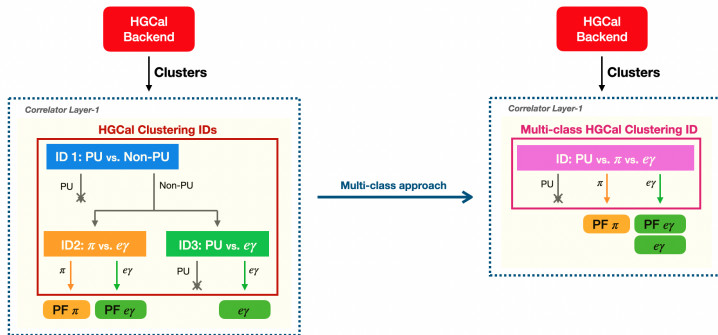
The Phase-2 HGCal will cover the pseudorapidity region from 1.52 to 3 with a high transverse and longitudinal granularity. HGCal trigger primitives with three-dimensional clustering information will be sent to the L1T Calorimeter and Correlator subsystems. Once the clusters arrive from HGCal backend electronics to the Correlator Trigger, the identification of the clusters takes place in the Correlator Layer-1. The current clustering ID system composed of three separate IDs is shown in the diagram on the left. The clusters that are not identified as pileup (PU) are required to pass two separate IDs for π and e/γ to build candidates for PF reconstruction and clusters for e/γ triggers. To harmonize and simplify the clustering ID design that is currently in place for Phase-2, a new clustering ID is developed with a multi-class approach using BDT.

HGCal Cluster Identification: Multi-class ID
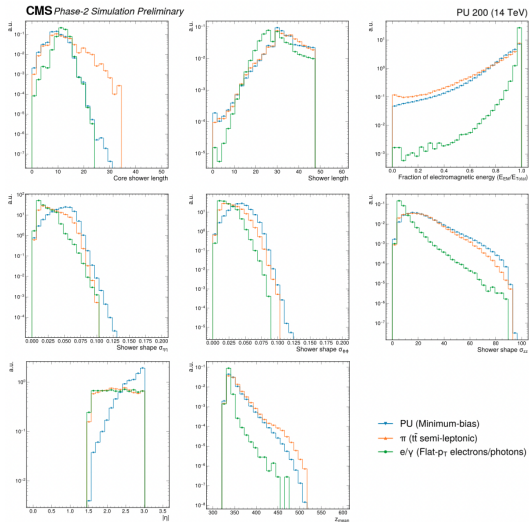
The plots show the distribution of the input features for building the multi-class BDT ID for each class. These input features based on HGCal cluster shower kinematics and shapes are used to train an XGBoost [4] BDT model. Training datasets are taken from Phase-2 simulation with 200 PU: minimum-bias samples for PU, $t\bar{t}$ semi-leptonic process for π, and flat p_T electron/photon samples for e/γ, after applying the cut of $p_T > 5$ GeV and $\eta < 2.4$ for all clusters and additionally requiring gen-matching of $\Delta R(\text{gen, reco}) < 0.1$ for π and e/γ clusters.
Eight cluster features are used as inputs, as shown in the plot from top left to bottom right: core shower length, shower length, fraction of electromagnetic energy E_{EM}/E_{Total}, shower shape $\sigma_{\eta\eta}$, $\sigma_{\phi\phi}$, σ_{zz}, absolute value of cluster pseudorapidity $|\eta|$, and z_{mean}. Shower lenghth measures the maximum span of the clusters in terms of the number of HGCal layers. Core shower length is the number of continuous shower layers. Shower shape $\sigma_{ii}(i = \eta, \phi, z)$ inputs refer to the energy-weighted RMS over cluster constituents. z_{mean} is calculated by the energy-weighted barycenter over the longitudinal coordinate.

HGCal Cluster Identification: Multi-class ID
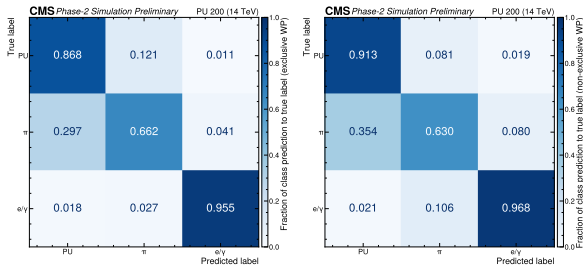
Multi-class BDT for the new clustering ID outputs a vector of BDT scores $\langle \text{PU}, \pi, e/\gamma \rangle$ given a cluster, where the BDT score for each class i represents the probability for the cluster to be of true label i. Given the BDT scores, Working Point (WP) is determined by evaluating on Phase-2 simulation with 200 PU: minimum-bias samples for PU, $t\bar{t}$ semi-leptonic process for π, and flat p_T electron/photon samples for e/γ. Two different schemes for the multi-class ID WP have been studied: *Exclusive WP* and *Non-exclusive WP*. *Exclusive WP* is to identify the cluster as the class with the maximum score from the vector of BDT class outputs. *Non-exclusive WP* is to define a cut on the BDT score for each class, allowing for a cluster to be sometimes identified as more than one class. This WP was optimized for the best efficiency and purity of e/γ clusters and to preserve the performance of the PF reconstruction of the MET.
Confusion matrices from using these WP schemes are displayed: *exclusive WP* (left) and *non-exclusive WP* (right). *Non-exclusive WP* scheme is chosen at the end for its better ID efficiency and downstream interpretability.
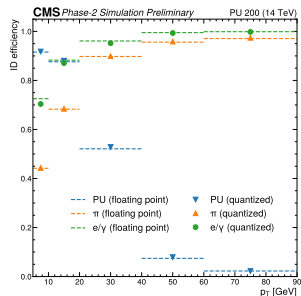
HGCal Cluster Identification: Multi-class ID

	BRAM	DSP	FF	LUT
SLR	0.0%	0.0%	1.0%	25.0%
Total	0.0%	0.0%	0.0%	6.0%

	Latency (cycles)	Latency (absolute)
Clock (2.78 ns)	5	13.890 ns

For hardware implementation of the model, the XGBoost model is synthesized via Conifer [3] library with Vitis HLS synthesis tools. The synthesis ran on the Xilinx Virtex UltraScale+ VU13P FPGA with a clock cycle of 2.78 ns. Floating value input features are quantized to ap_fixed $\langle 20, 10 \rangle$, corresponding to seven bits for the integer and ten bits for the decimal. The tables show the latency and FPGA resources used by the new multi-class clustering ID in terms of the used fraction of Block Random Access Memory (BRAM), Digital Signal Processors (DSPs), flip-flops (FF), and lookup tables (LUTs), both in the Super Logic Region (SLR) and the whole FPGA.
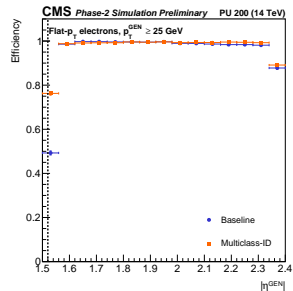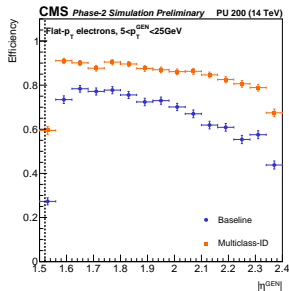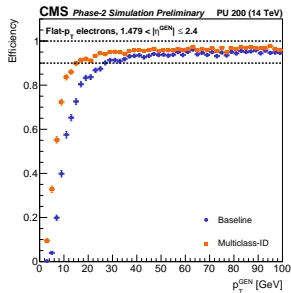
The plot shows the ID efficiency for each class as a function of p_T, comparing between XGBoost BDT model with floating inputs and the synthesized model with quantized inputs. The evaluation of the ID efficiency is done on Phase-2 simulation with 200 PU: minimum-bias samples for PU, $t\bar{t}$ semi-leptonic process for π, and flat p_T electron/photon samples for e/γ clusters.

# HGCal Cluster Identification: Multi-class ID

The plots compare the efficiency of reconstructing a genuine electron at L1T for the baseline and the new multiclass e/γ IDs for standalone electromagnetic objects on a flat-p_T electron sample. The efficiency as a function of the generator-level p_T (left) and pseudorapidity (middle and right) is illustrated. In the central and right plots, the vertical dashed line represents the transition region between the barrel and the endcaps. The efficiency at the low p_T region is significantly improved with the new multi-class ID, e.g. an overall 35% improvement in efficiency for $5 < p_T < 10$ GeV. The efficiency improves throughout the whole endcap η coverage. While only high-p_T standalone electromagnetic objects ($p_T > 25$ GeV) are used in the TDR Trigger Menu [1], efficient identification of low-p_T clusters can be exploited in building track-matched electrons and in the L1T Scouting system to target signals with low-p_T electrons.

HGCal Cluster Identification: Multi-class ID
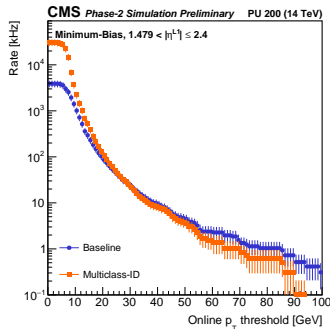
The plot shows the rates of standalone electromagnetic objects measured on minimum bias samples, comparing the current baseline and new multi-class clustering ID. The multi-class ID increases the rate in low p_T regions but lowers it at the typical thresholds used in the TDR baseline Trigger Menu ($p_T > 25\,\text{GeV}$).

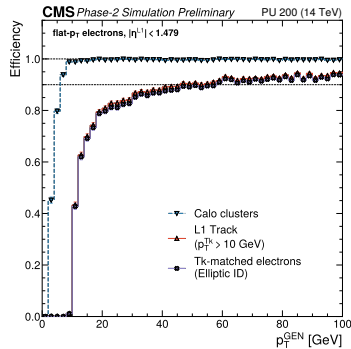# Electrons in the Phase-2 Level-1 Trigger in the ECAL barrel region
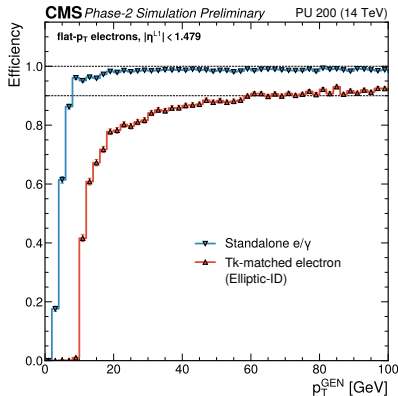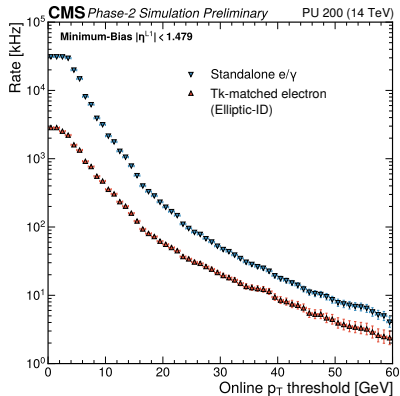
Matching calorimeter deposits to tracks for identifying electrons is pivotal for sustaining the typical Single and Double Electron trigger thresholds of the TDR Menu. The plot on the left illustrates the comparison between the rate of events on a Minimum Bias sample with an average of 200 pileup when requiring a standalone calorimeter object or a track-matched object at different p_T thresholds. The right plot shows the trigger efficiency as a function of the transverse momentum of the generated electron for the same objects. The efficiency plots are obtained on a flat-p_T electron sample using the TDR baseline matching algorithm called *Elliptic-ID* [1].
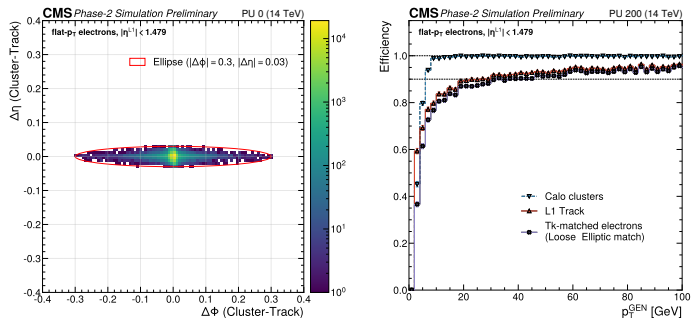
Track-matched Electron Identification in the barrel

The baseline identification technique (*Elliptic-ID*) relies on the independent selection of electromagnetic calorimeter clusters and L1 tracks followed by an angular matching procedure [1]. Electrons' track quality and momentum resolution often suffer because of distortions due to bremsstrahlung radiation. A minimum p_T cut of 10 GeV is applied to the track to control the combinatorics and the rate.

A new machine learning approach called *Composite-ID*, combines information about tracks and clusters in the barrel into a single model for matching and identification.

CMS *Phase-2 Simulation Preliminary*   PU 200 (14 TeV)

flat-p_T electrons, $|\eta^{L1}| < 1.479$

- Calo clusters
- L1 Track ($p_T^{Tk} > 10$ GeV)
- Tk-matched electrons (Elliptic ID)

p_T^{GEN} [GeV]

# Tk-matched Electrons in the barrel: Composite-ID Candidates

Candidate track-calorimeter cluster pairs are selected with a loose elliptic match ($\Delta\phi < 0.3$, $\Delta\eta < 0.03$) and no requirements on the track and the cluster.

For training the model, signal candidates are selected based on the geometrical ($\Delta R < 0.1$) matching of the cluster with a generated electron particle, while all candidates in a minimum-bias sample are used for the background.

The left plot shows the Track-Cluster $\Delta\eta$ vs $\Delta\phi$ 2D distribution in a 0 PU sample, while the right plot shows the efficiencies of the standalone object, L1 tracks, and Tk-matched candidates as a function of the generated electron p_T.

Tk-matched Electrons in the barrel: Composite-ID Model
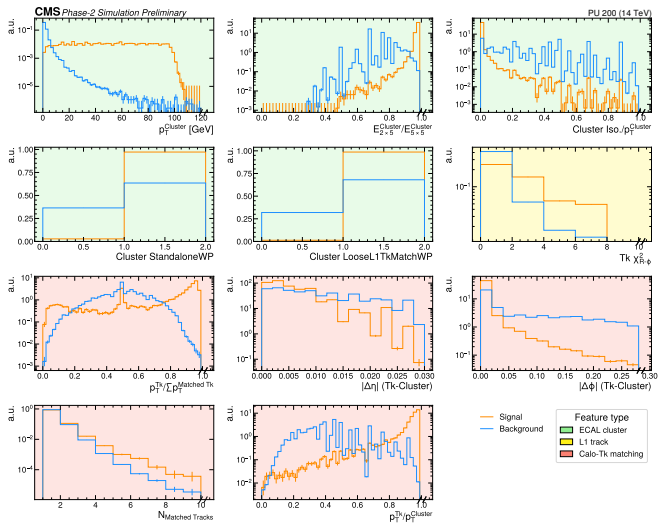
The Composite-ID combines calorimeter cluster shape variables, track qualities, and track-matching features in a single BDT model controlling the identification of track and calorimeter deposit and the tightness of the matching.

The plots show the distributions of the input features for signal and background, where $E_{2\times5}/E_{5\times5}$ is the ratio of energy in a 2×5 to a 5×5 ECAL crystal array, Tk $\chi^2_{R-\phi}$ is the track fit χ^2 in the $R - \phi$ plane, while StandaloneWP and LooseL1TkMatchWP are cluster quality flags. The signal is defined as track-matched candidates matched to a generated electron in a flat-p_T electrons sample, while all track-matched candidates in a Minimum-bias sample are used for the background.

Tk-matched Electrons in the barrel: Composite-ID Model
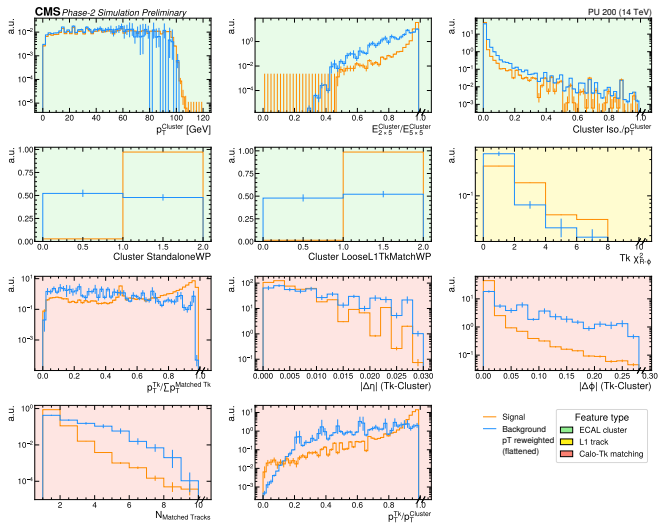

To prevent the model from imposing a tight cut on $p_T^{Cluster}$, the features of all the background candidates used for training were reweighted to flatten the $p_T^{Cluster}$ distribution (top left plot). This improves the efficiency for low-p_T electrons. The plots show the input features of the model after reweighting.

Unlike the training, the validation is done using the non-reweighted features.

Tk-matched Electrons in the barrel: Composite-ID Model

The BDT scores for signal and background are obtained from the bit-wise emulator implemented using Conifer with `ap_fixed` precision. The plot on the right shows the receiver operating characteristic curve (ROC) curve of the BDT model *and the area under the curve (AUC)* using the bit-wise firmware emulator in different bins of p_T^{Cluster}.
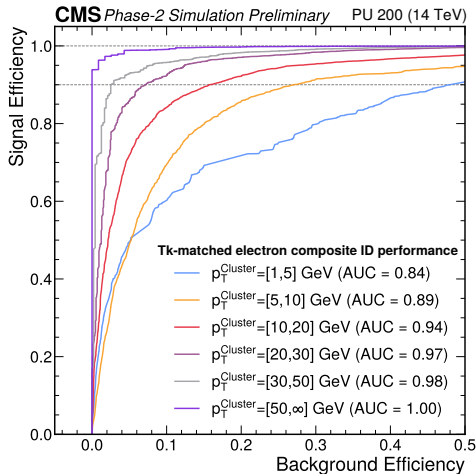
Tk-matched Electrons in the barrel: Elliptic vs Composite ID performance
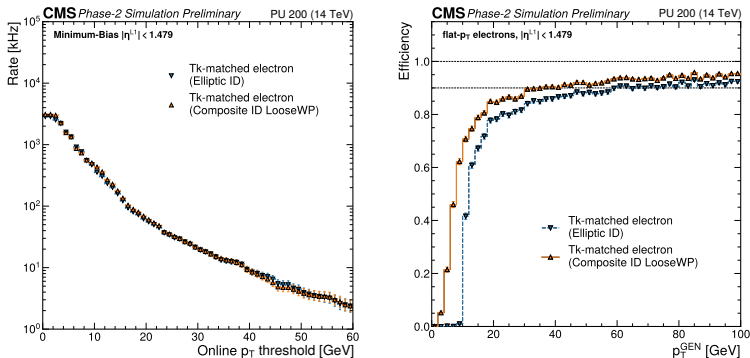
To compare the Composite-ID with the Elliptic-ID, two working points (WPs) applied in different p_T bins were created: a tight WP that matches the efficiency of the Elliptic-ID and a loose WP that matches the rate of the elliptic-ID.

The left plot compares the rate as a function of the cluster p_T, while the right plot shows the efficiency as a function of the p_T^{Gen}. The new Composite-ID model allows for a substantial efficiency gain compared to the elliptic-ID at parity of rate (Loose WP), especially in the low p_T region. This is very beneficial for the typical Double Electron seed of the baseline Menu, which has a threshold of 25, 12 GeV.

# Tk-matched Electrons in the barrel: Elliptic vs Composite ID performance
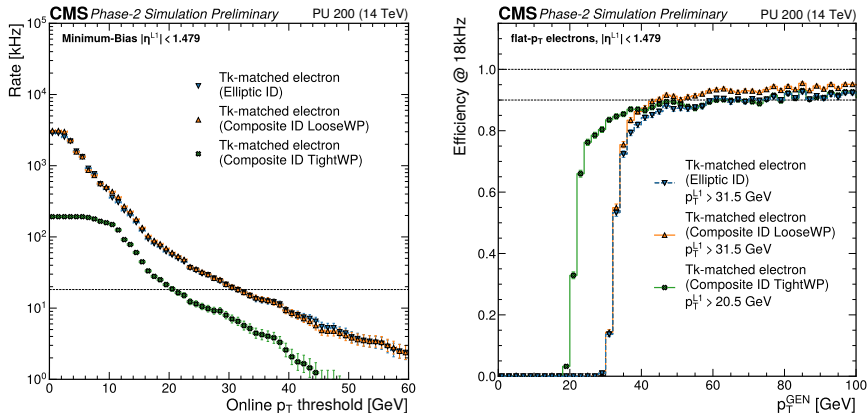
Single lepton trigger performance using the Elliptic-ID and the Composite-ID with different working points. The left plot illustrates the rate as a function of the cluster p_T. The right plot compares the efficiency as a function of the p_T of the generated electron for a fixed 18 kHz rate, corresponding to the barrel bandwidth for Run-3 like thresholds.

Tk-matched Electrons in the barrel: Composite-ID firmware emulation

The model has been synthesized in firmware using the Conifer [3] library and Vivado HLS. The input features are quantized using the ap_fixed format with nine bits for the integer and fifteen bits for the decimal precision, while the raw scores are quantized using 4 and 12 bits for the integer and decimal parts, respectively. The model is synthesized for a frequency of 180MHz, matching the implementation of the egamma reconstruction in the Correlator Layer-1 boards. The table illustrates the estimated resource utilization on a Xilinx Virtex UltraScale+ VU13P FPGA. The plot on the right shows the ROC curves for the XGBoost [4] model and the quantized model emulated with Conifer.

	BRAM	DSP	FF	LUT
SLR	0.0%	0.0%	0.12%	6.5%
Total	0.0%	0.0%	0.02%	1.6%

	Clock cycles	Latency
Clock (5.56 ns)	5	27.8 ns
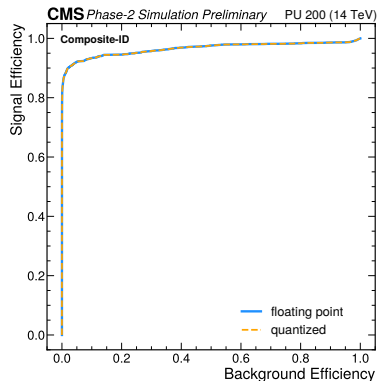
References

📄 CMS Collaboration , "The Phase-2 Upgrade of the CMS Level-1 Trigger", *CERN-LHCC-2020-004* https://cds.cern.ch/record/2714892?ln=en.

📄 CMS Collaboration , "Electron Reconstruction and Identification in the CMS Phase-2 Level-1 Trigger", *CERN-CMS-DP-2023-047* https://cds.cern.ch/record/2868782?ln=en.

📄 Conifer, https://github.com/thesps/conifer.

📄 XGBoost, https://xgboost.readthedocs.io/en/stable/.