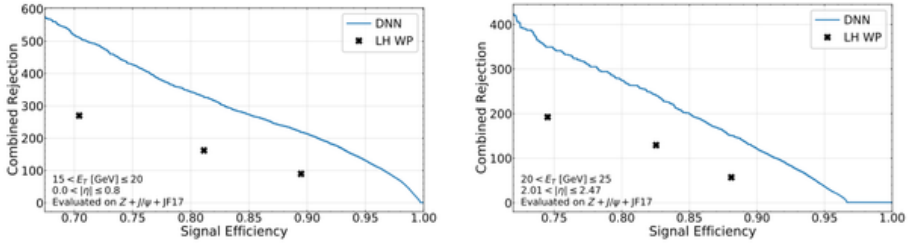
Table of Contents

DNN electron identification

A new electron identification algorithm based on a multiclass deep neural network (DNN) is provided on an experimental basis (i.e. **for testing, not to be used in publications yet**). It relies upon the same discriminant variables as the likelihood-based identification as inputs, but provides enhanced rejection:



More information is available in this presentation, as well as in ATL-COM-PHYS-2020-829.

Usage

Three working points are predefined, with similar efficiencies as their likelihood equivalent. Precalculated decision flags are available as decorations in derivations starting with release 21.2.172:

`DFCommonElectronsDNNLoose`, `DFCommonElectronsDNNMedium`, `DFCommonElectronsDNNTight`

In addition, the raw classifier scores are made available as well for the signal and five background categories:

`DFCommonElectronsDNN_pel`, `DFCommonElectronsDNN_pcf`, `DFCommonElectronsDNN_ppc`, `DFCommonElectronsDNN_phf`, `DFCommonElectronsDNN_ple`, `DFCommonElectronsDNN_plh`

Experimental scale factors in release 21.2 are available for the three predefined identification working points, isolation and a choice of triggers, and can be obtained as usual with the AsgElectronEfficiencyCorrection tool; the calibration files to employ in this case are the following:

Type	Path to calibration file	details
Identification	path to file	in progress
Isolation	path to file	in progress
Trigger	path to file	in progress

Unlike likelihood-based ID, the DNN ID decisions are currently calculated after applying shower shape corrections to the simulation, which mostly explains why the scale factors differ so much.

Interplay with trigger

Run 2 electron trigger chains relied upon likelihood-based identification. Using such a trigger in combination with an offline DNN working point can induce extra inefficiencies at the level of up to 10%, as the two identification algorithms do not always select the same signal events. Possible solutions have not been studied much yet, but may be among these:

- using a looser online working point than offline (e.g. DNNTight offline with HLT_2e17_lhvloose_nod0_L12EM15VHI)
- using a tighter online working point than offline (e.g. DNNMedium offline with HLT_e26_lhtight_nod0_ivarloose), in case this still brings additional rejection.
- using likelihood offline (+ trigger matching) only for the number of electrons requested by the trigger, and DNN for the additional electrons required in the final state

etc

For Run 3, trigger chains relying on DNN identification are being developped and will allow to avoid these limitations.

Feedback request

The Egamma group encourages analyses to try this new identification tool, which will play a major role in Run 3, and provide feedback about it. Several points of attention are:

- data/MC agreement in various kinematic regions, after applying scale factors
- suitability of the predefined working points
- achieved background rejection in the analysis' phase space of interest, compared to the likelihood ID
- proposal of new working points (which can be tried by playing with the raw classifier scores)
- retained strategy to cope with online trigger identification
- behaviour of the DNN ID for signal events with unusual environments (e.g. boosted tops)

etc

This topic: Sandbox > JulienMaurerSandbox
Topic revision: r20 - 2021-10-11 - JulienMaurer