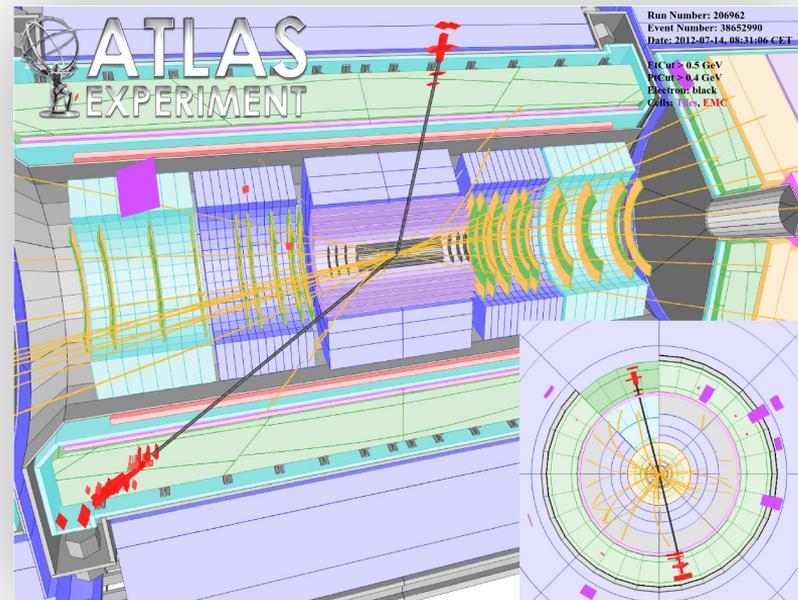


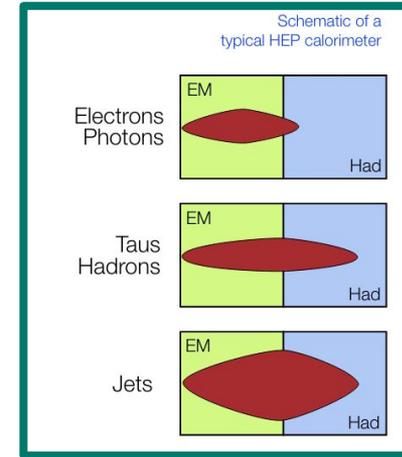
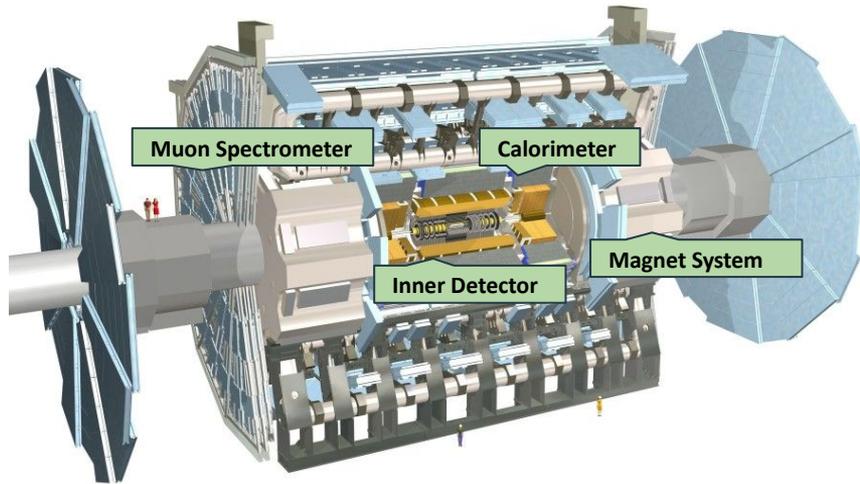
Offline electron identification (ID) in ATLAS using a Deep Neural Network

Jornadas científicas IFIC:
Origin of mass. L1 Higgs Force

Enrique Valiente Moreno (IFIC, CSIC-UV),
on behalf of the ATLAS Collaboration

02/12/2024





- **Electron identification** is an essential input to many ATLAS physics analyses involving the **Higgs boson, the W/Z bosons and top quarks**, as well as phenomena beyond the Standard Model (BSM)
 - Clean electromagnetic signature in the detector
 - Different calorimeter activity than other object
- Mostly **prompt electron** from W and Z bosons, τ leptons or BSM particles **are relevant**
 - Other sources considered as **background**, including jets mimicking their signature

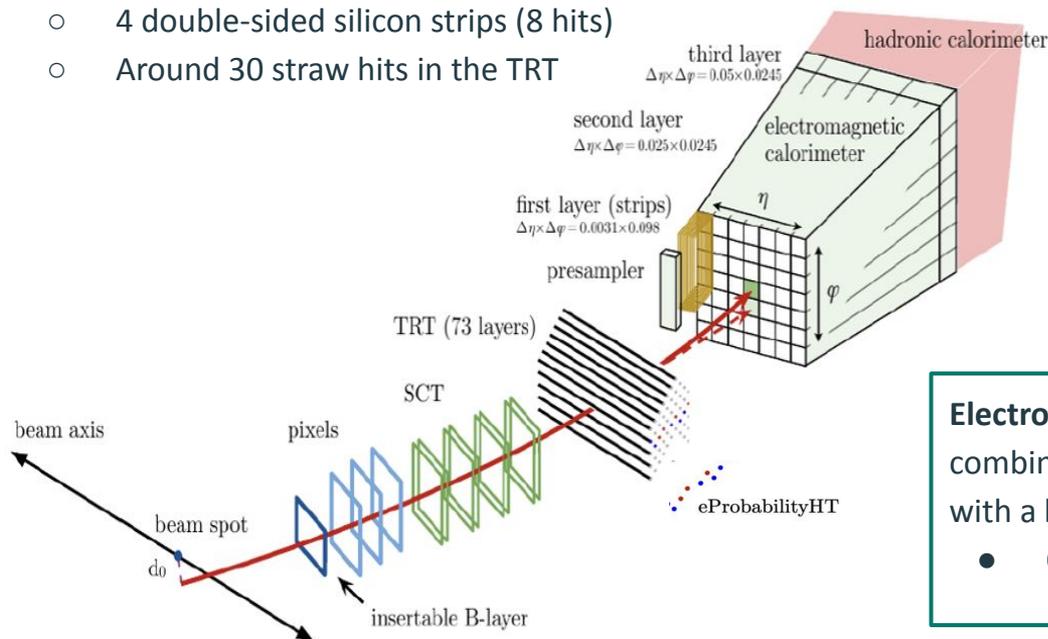
Particles produced in the collision leave different signatures at each layer of ATLAS detector

1. Inner detector records the tracks of charged particles and transverse momentum

- An electron usually hits:
 - 3 pixel layers (+ insertable B-layer)
 - 4 double-sided silicon strips (8 hits)
 - Around 30 straw hits in the TRT

2. Calorimeters register the particle energy and position after its decay into a shower

- Commonly deposits its energy in four successive EM calorimeter layers:
 - pre-sampler (early energy loss)
 - 1st high-granularity η strips layer
 - 2nd layer (collects most of the energy)
 - backplane layer (tails of the shower)
 - Small amount of electron's energy left in the hadronic calorimeter



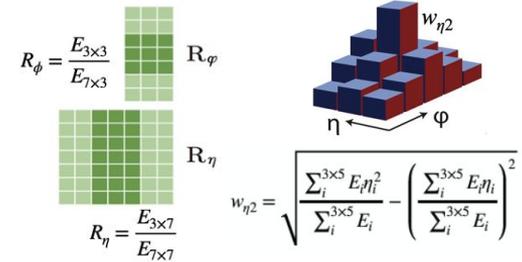
Electron candidates reconstructed and identified as a combination of a **charged track in the inner detector** with a **localised energy cluster in the calorimeter**

- Close matching in $\eta\chi\phi$ space

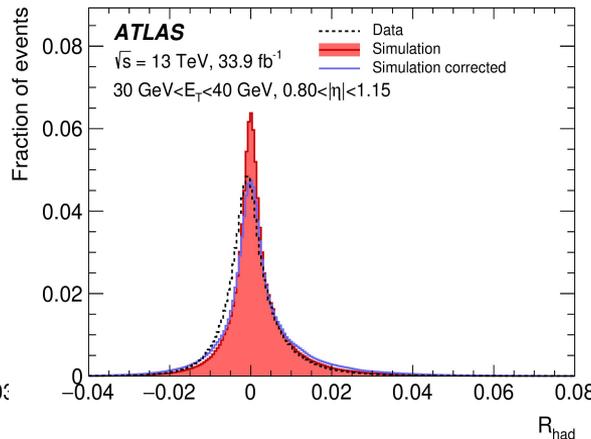
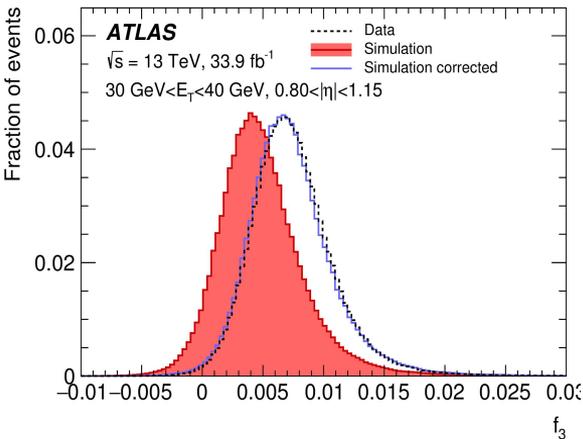
- **ATLAS performed electron ID via a Likelihood approach**
 - separate **prompt, isolated** electrons from objects with similar signatures
- **Likelihood of being signal or background obtained as product of input variables PDFs**
 - **Binomial discriminant** defined as the likelihoods ratio
- Highly discriminating input variables build out of the objects signatures in the detector:
 - Track properties
 - Track-to-cluster matching variables
 - Shape properties of electromagnetic shower in the calorimeter (shower-shapes)

$$L_{S(B)}(\mathbf{x}) = \prod_{i=1}^n P_{S(B),i}(x_i)$$

$$d_L = \frac{L_S}{L_S + L_B}$$



[Eur. Phys. J. C 79 \(2019\) 639](#)



Modelling

Biggest concern to use MC to train ML/AI architectures are **differences** with respect to Data

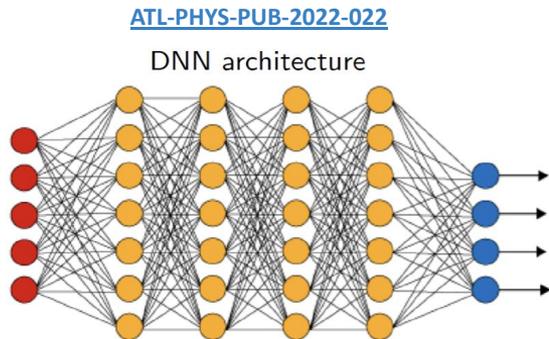
- Uncertainty in the detector simulation...
 - **Affect shower shape**
 - Data/MC disagreements on efficiency performance
 - **Usual solution:** independent local correction (**shower-shape fudging**)

Electron identification (ID) using a Deep Neural Network

- **Deep Neural Networks (DNN)** are powerful for **classification** and signal to background **discrimination**
- DNN exploits **correlations** between inputs, leading to a more performant algorithm

DNN architecture:

- 5 **hidden layers** with 256 nodes each
- Activation function in each layer: **leaky ReLU**
- **Batch normalization** is also used



- Trained using Monte Carlo simulated electrons
 - **Shower-shape** related variables **corrected** to match Data
 - **Similar input variables** as used for the LH discriminant

Output layer:

- Six outputs: signal + 5 background classes, with **softmax** activations for **multiclass** classification

Class	Description
Prompt Electron (EI)	Prompt isolated electrons coming from Z, W and J/Psi
Charge Flip (CF)	Same as prompt, but with incorrectly reconstructed charge
Photon Conversion (PC)	Electrons from prompt photons
Heavy-Flavour (HF)	Electrons coming from a b- or c- hadron decay
Light-Flavour e/ γ	Electron coming from a u-, d- or s- hadron
Light-Flavour Hadrons	Undecayed hadrons

Flexibility → different discriminants can be defined to enhance certain background rejection combining the 6 DNN scores

Prompt vs. background

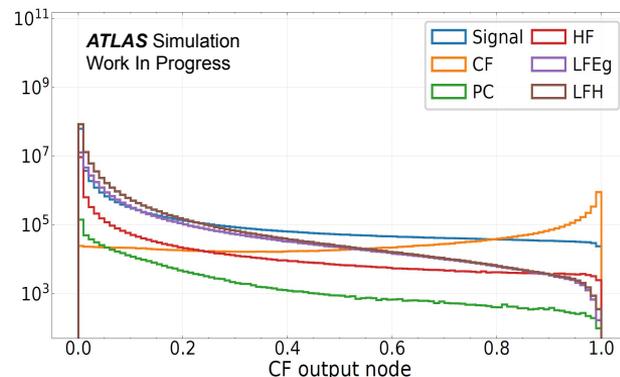
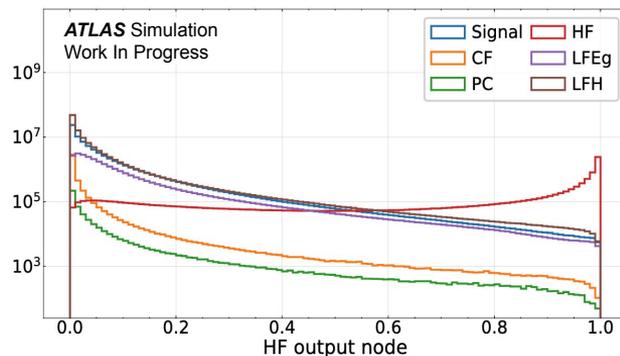
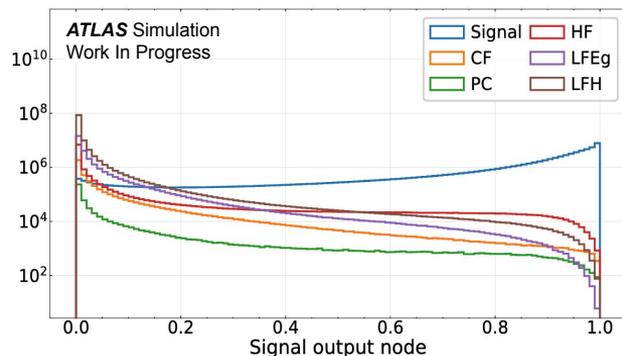
$$D_{el} = \ln \left(\frac{f_{El} p_{El} + (1 - f_{El}) p_{CF}}{f_{PC} p_{PC} + f_{HF} p_{HF} + f_{LFEg} p_{LFEg} + (1 - f_{PC} - f_{HF} - f_{LFEg}) p_{LFH}} \right)$$

Signal vs. CF

$$D_{CF} = \frac{p_{El}}{p_{CF}}$$

- p_x : DNN score for each class
- f_x : relative fraction from each class

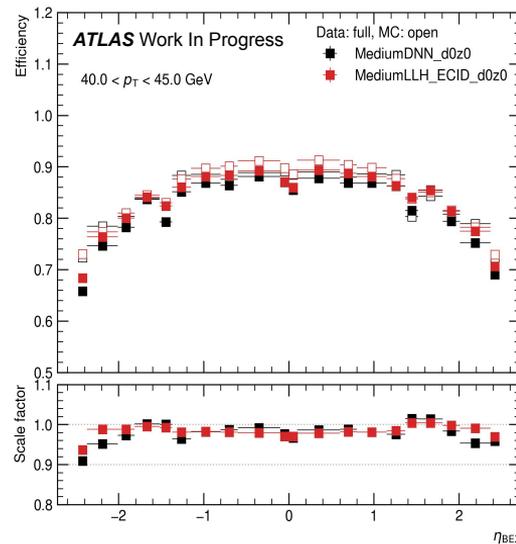
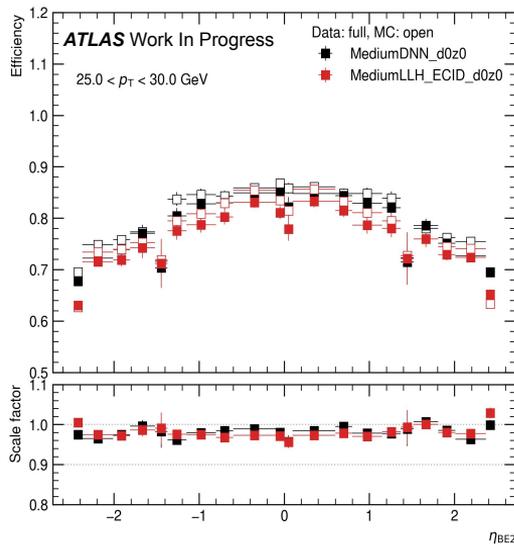
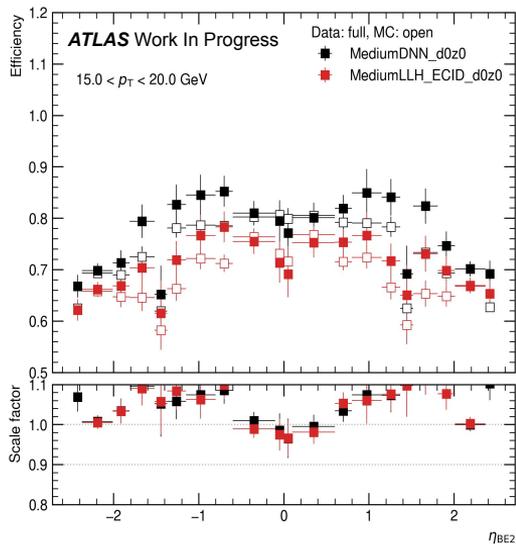
- Outstanding discrimination can be achieved not only for **signal vs. background**, but also **for the different background classes against the others**



- Working points (WPs) can be obtained defining cuts on discriminants targeting pre-defined signal ID efficiencies

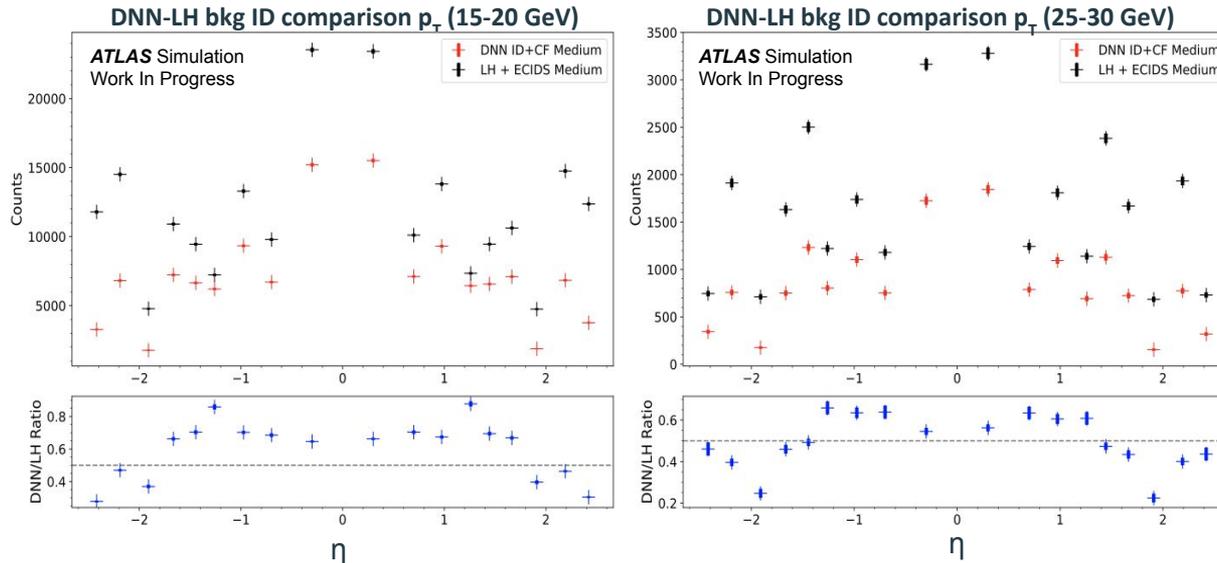
Efficiencies estimated directly from data using tag-and-probe methods:

- Select unbiased samples of prompt electrons from well known decays ($Z \rightarrow e+e^-$ or $J/\psi \rightarrow e+e^-$)
- One of the electrons must satisfy strict selection requirements (**tag**), the other very loose ones (**probe**)
- Efficiency computed by applying selections on the probe sample (after subtracting any remaining background)



Overall background rejection:

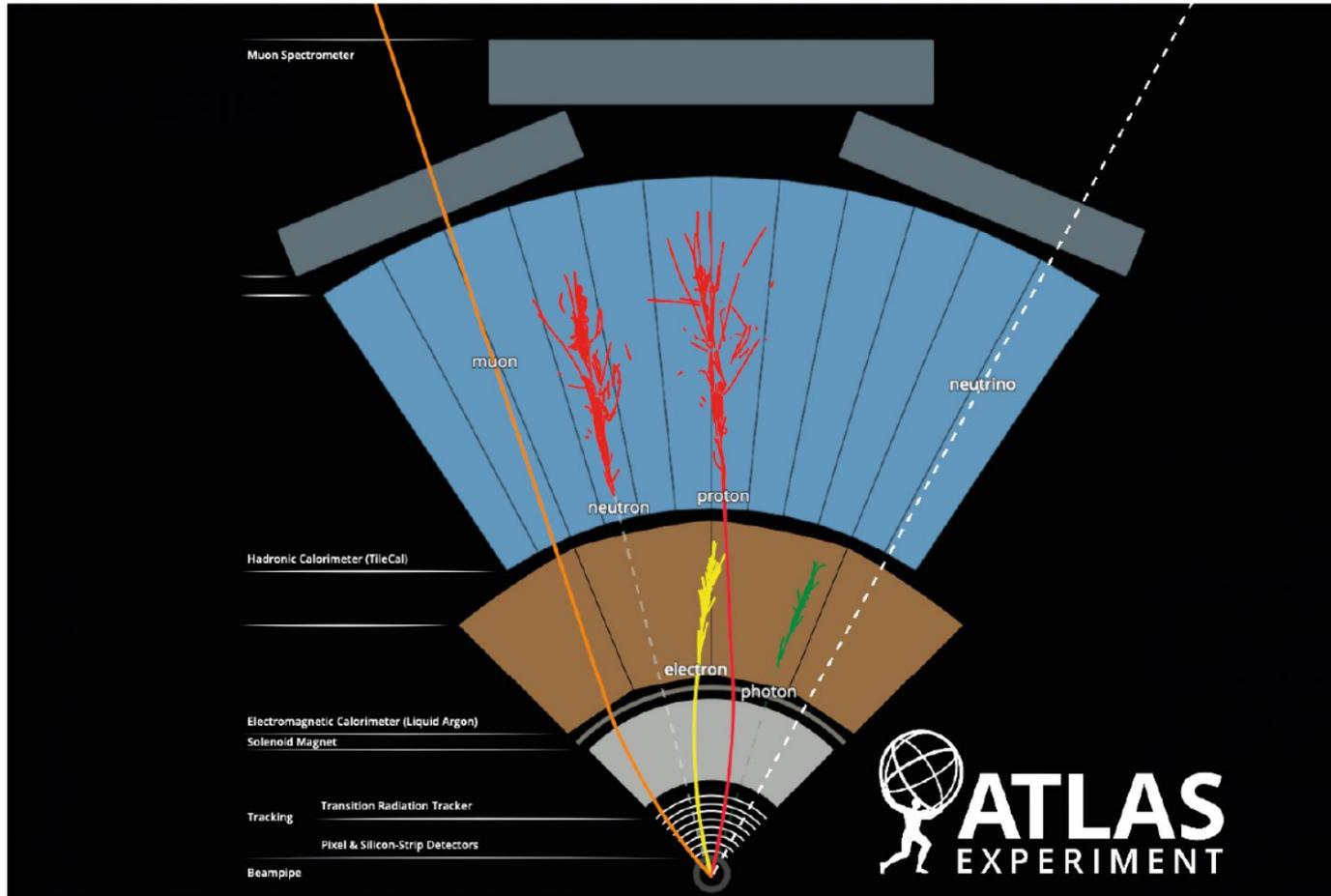
- Evaluated on background MC sample with all tree-level 2 \rightarrow 2 QCD processes + top/W/Z production (mix of background classes, dominated by QCD jets)
- Expected improvement of **DNN** over **LH** is not mainly due to the use of a **more complex architecture**, but to the way it “cuts” on input variables
 - Includes some extra variables (devoted to improve CF electrons classification)
 - Benefits by learning from correlations



- New ML algorithm implemented for electron ID with multiple advantages
 - Multiscore classification performed allowed for different electron classes (signals and backgrounds)
 - DNN profits from correlations between input variable distributions, unlike LH
 - Flexibility: wide variety of discriminant could be defined to enhance certain background rejection and be optimized for the electron composition of each phase space (thanks to the f_x)
- Data/MC agreement in shower-shapes is one of the most issues we need to face
 - Cutting-edge ML/AI methods are being also studied in order to tackle this problem in a general way
- This DNN algorithm has been already tested in some particular analysis cases (H→4l) showing good performance
 - Already implemented in most of our datasets to be tested by analysis team
- Next steps: continue with its implementation and improvement for the new ATLAS Run-3 data (2022-2024)

- The author's work is supported by:
 - FPU2021 grant funded by MICIU/AEI /10.13039/501100011033 and by the FSE+
 - Grants PID2021-124912NB-I00 and PID2021-125069OB-100 funded by MCIN/AEI/10.13039/501100011033
 - Project ASFAE/2022/008 funded by MCIN, by the European Union NextGenerationEU (PRTR-C17.I01) and Generalitat Valenciana

Additional material



Monte Carlo samples used to generate input n-tuples for the DNN

MC20(e) SAMPLES

Zee

mc16_13TeV:mc16_13TeV.361106.PowhegPythia8EvtGen_AZNLOCTEQ6L1_Zee.deriv.DAOD_EGAM1.e3601_e5984_s3126_r10724_r10726_p4615

JF17

mc20_13TeV.423300.Pythia8EvtGen_A14NNPDF23LO_perf_JF17.deriv.DAOD_EGAM7.e3848_s3681_r13145_p4940

ttbar

mc20_13TeV.410470.PhPy8EG_A14_ttbar_hdamp258p75_nonallhad.deriv.DAOD_EGAM1.e6337_s3681_r13145_p4940

Jpsi

mc20_13TeV:mc20_13TeV.423210.Pythia8B_A14_CTEQ6L1_bb_Jpsie3e3.deriv.DAOD_PHYS.e4364_s3681_r13145_r13146_p5057

Sample ID	Process	Events
005802	JF17 (rel 13)	1.9M
105802	JF17 (rel 14)	2.4M
105012	J3 (rel 14)	396k
105200	T1 (rel 14)	467k
105403	SU3 (rel 14)	10k
106400	SU4 (rel 14)	49k

JF17 sample

- ▶ Generation: QCD hard scattering, direct photons, W, Z, top, etc.
- ▶ Filter: More than 17 GeV in a $\Delta\phi \times \Delta\eta = 0.12 \times 0.12$ region

Filters a lot of energy in a very narrow region of the EM calorimeter: jets which could mimic the electron signature

- All 2→2 scattering processes
- Mostly QCD due to XS

Flexibility → different discriminants can be defined to enhance certain background rejection combining the 6 DNN scores

Prompt vs. background

$$D_{el} = \ln \left(\frac{f_{El} P_{El} + (1 - f_{El}) P_{CF}}{f_{PC} P_{PC} + f_{HF} P_{HF} + f_{LFEg} P_{LFEg} + (1 - f_{PC} - f_{HF} - f_{LFEg}) P_{LFH}} \right)$$

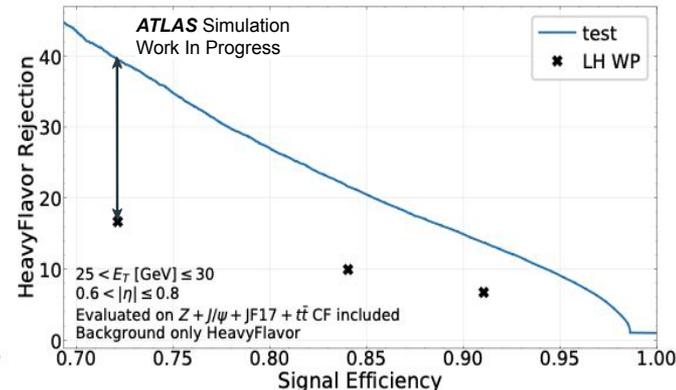
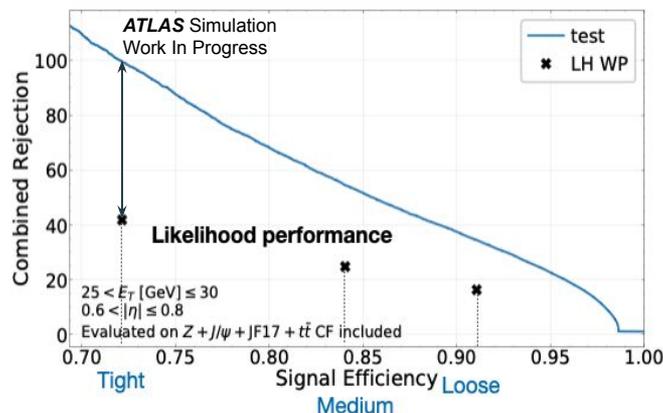
Signal vs. CF

$$D_{CF} = \frac{P_{El}}{P_{CF}}$$

- p_x : DNN score for each class
- f_x : relative fraction from each class

- Discriminating performance is obtained by comparing signal efficiency (ϵ) and background rejection ($1/\epsilon$) at different probability cuts

- At ~70% of signal efficiency, **DNN outperforms LH** by more than a factor of two



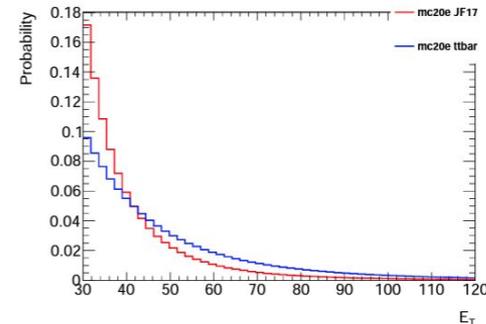
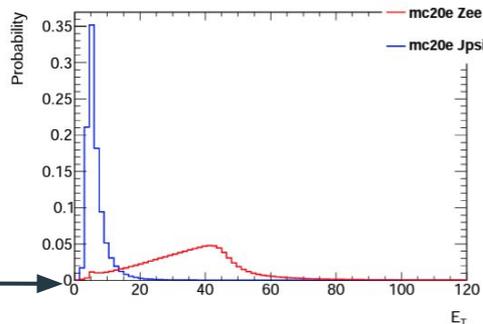
- Working points (WPs) can be obtained defining cuts on discriminants targeting pre-defined signal ID efficiencies
 - Electron ID WPs tuning performed in both p_T and η bins

- For the multinomial classification we consider different types of background electrons:
 - HF:** electrons coming from heavy-flavour quark decays
 - PC:** prompt photon convert into $e+e^-$ pair
 - LFH:** light-flavour jets can also mimic the prompt electron signatures
 - LFEg:** electrons from the decay of u^- , d^- , s^- hadrons.
- Charge flipped (CF)** electrons can be considered as signal (similar to LH). However **Signal** only refers to prompt electrons.
- Selection of candidates from different samples purely based on simulation information (at MC truth level), using IFF, TruthType (TT), and TruthOrigin:

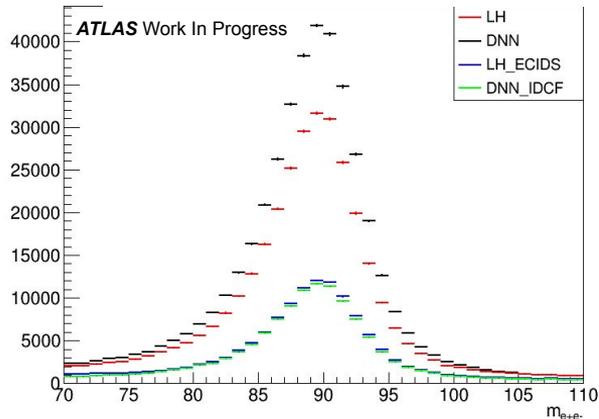
Class	Definition	
Signal	IFF = 2	} → Zee & Jpsi
ChargeFlip	IFF = 3	
PhotonConv	IFF = 5	} → ttbar
HeavyFlavor	IFF = 8 IFF = 9	
LFEgamma	IFF = 10 & (TT = 4 TT = 16)	
LFHadron	IFF = 10 & TT = 17	} → JF17
Others	everything else	

- Z** → e^+e^- and **JF17** are the main signal and background sources
- Jpsi** could be used to increase signal and CF statistics at low E_T . Same for **ttbar** for background at higher values.

Offline electron E_T distributions



DNN-LH Same Sign invariant mass



CF rejection: comparison between DNN and LH

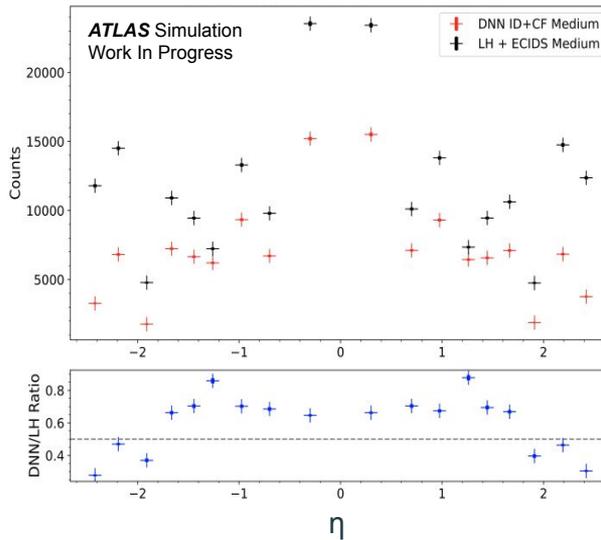
- $Z \rightarrow e+e-$ data events in which one of the electrons has its charge mis-identified
 - m_{e-e+} distributions for SS data events
 - Lower peak as possible around Z-mass means powerful CF background rejection

Overall background rejection:

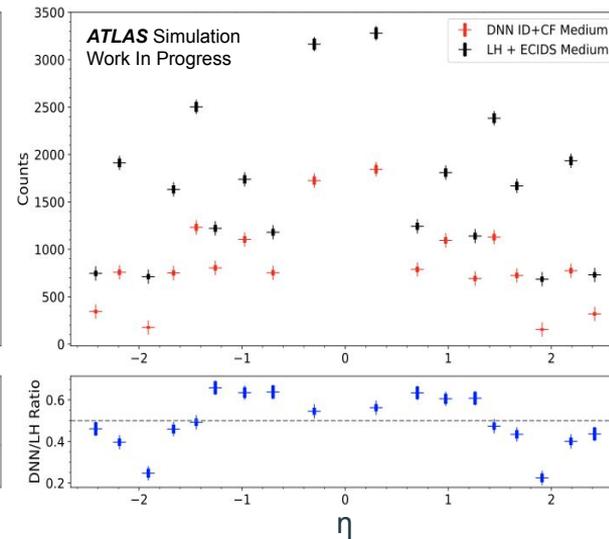
(evaluated on Monte Carlo JF17 background)

- Expected improvement of **DNN** over **LH** is not mainly due to the use of a **more complex architecture**, but to the way it “cuts” on input variables
 - Includes some extra variables
 - Benefits by learning from correlations

DNN-LH bkg ID comparison p_T (15-20 GeV)

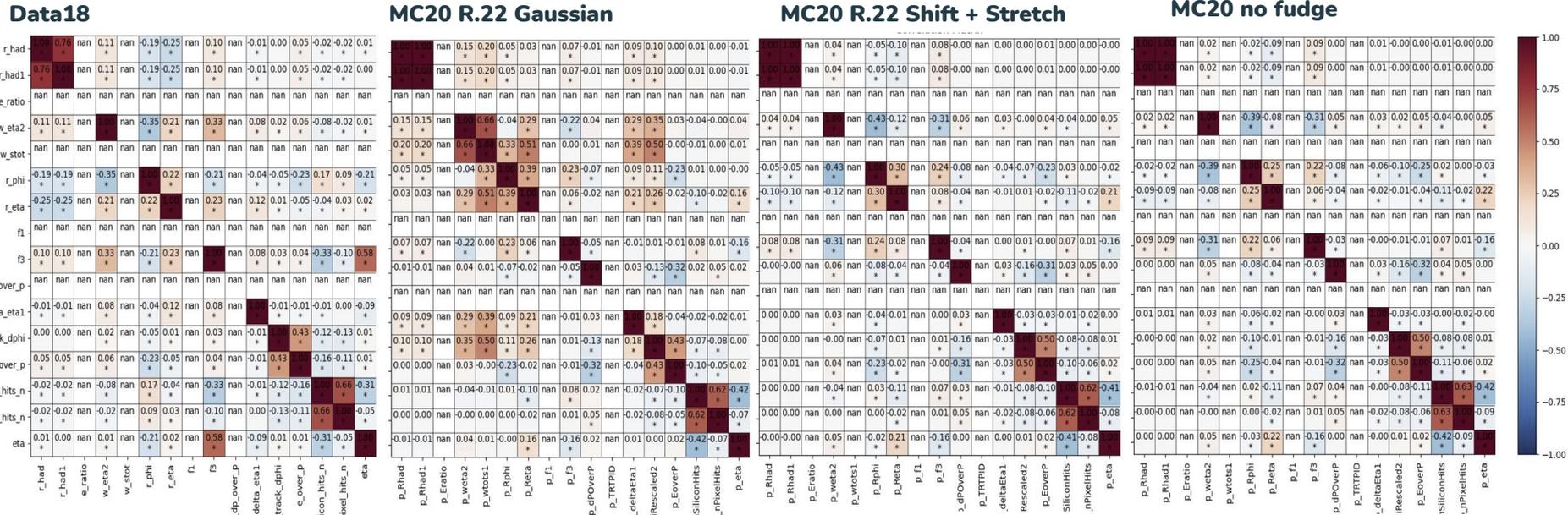


DNN-LH bkg ID comparison p_T (25-30 GeV)



- **Correlations between variables:** key ingredient in ML/AI techniques
 - Not perfectly modelled either in MC
 - Can be broken by certain correction approaches

Correlations between input variables for Electron ID DNN



Class	Description	Label	Sample
Prompt Electrons	Prompt isolated electrons, e.g. electrons from $Z \rightarrow ee$, $W \rightarrow ev$, $J/\psi \rightarrow ee$ decays with the J/ψ being produced in the hard scatter. Electrons from a final state radiation photon or bremsstrahlung are also considered here if the origin is a prompt electron. Furthermore, the reconstructed charge is the same as the true charge.	El	$Z \rightarrow ee$, $J/\psi \rightarrow ee$
Charge Flip	Same as prompt electrons, but the reconstructed charge is the opposite of the true charge. In case of an electron originating from bremsstrahlung, the charge of the original prompt electron is considered the true charge.	CF	$Z \rightarrow ee$, $J/\psi \rightarrow ee$
Photon Conversion	Electrons from prompt photons which convert into an e^+e^- pair. Prompt photons which are reconstructed as an electron are also considered for this class.	PC	JF17, $t\bar{t}$, $Z\gamma$
Heavy-Flavour	Electrons coming from a decay of a b^- or c^- hadron. Prompt quarkonium decays such as $J/\psi \rightarrow ee$ where the J/ψ is produced in the hard scatter event are not included here but rather as prompt electrons.	HF	JF17, $t\bar{t}$
Light-Flavour e/γ	Electrons and photons from a decay of a u^- , d^- , or s^- hadron. This also includes for instance electrons from intermediate photon conversions: $\pi^0 \rightarrow \gamma\gamma$ with $\gamma \rightarrow ee$.	LFeg	JF17
Light-Flavour Hadrons	Undecayed hadrons	LFH	JF17

Type	Description	Name	Usage
Hadronic leakage	Ratio of E_T in the first layer of the hadronic calorimeter to E_T of the EM cluster	R_{had1}	DNN
	Ratio of E_T in the hadronic calorimeter to E_T of the EM cluster	R_{had}	DNN
Third layer of EM calorimeter	Ratio of the energy in the third layer to the total energy in the EM calorimeter. Due to known mismodelling at high $ \eta $, this variable is set to a default value for candidates with $ \eta > 2.01$	f_3	DNN
Second layer of EM calorimeter	Lateral shower width, $\sqrt{(\sum E_i \eta_i^2) / (\sum E_i) - ((\sum E_i \eta_i) / (\sum E_i))^2}$, where E_i is the energy and η_i is the pseudorapidity of cell i and the sum is calculated within a window of 3×5 cells	$w_{\eta 2}$	DNN
	Ratio of the energy in 3×3 cells over the energy in 3×7 cells centred at the electron cluster position	R_ϕ	DNN
	Ratio of the energy in 3×7 cells over the energy in 7×7 cells centred at the electron cluster position	R_η	DNN
First layer of EM calorimeter	Shower width, $\sqrt{(\sum E_i (i - i_{max})^2) / (\sum E_i)}$, where i runs over all strips in a window of $\Delta\eta \times \Delta\phi \approx 0.0625 \times 0.2$, corresponding typically to 20 strips in η , and i_{max} is the index of the highest-energy strip	w_{stot}	DNN
	Ratio of the energy difference between the maximum energy deposit and the energy deposit in a secondary maximum in the cluster to the sum of these energies	E_{ratio}	DNN
	Ratio of the energy in the first layer to the total energy in the EM calorimeter	f_1	DNN
Track conditions	Number of hits in the innermost pixel layer	n_{Blayer}	C
	Number of hits in the pixel detector	n_{pixel}	DNN/C
	Total number of hits in the pixel and SCT detectors	n_{Si}	DNN/C
	Transverse impact parameter relative to the beam-line	d_0	DNN
	Significance of transverse impact parameter defined as the ratio of d_0 to its uncertainty	$ d_0 /\sigma(d_0)$	DNN
	Momentum lost by the track between the perigee and the last measurement point divided by the momentum at perigee	$\Delta p/p$	DNN
TRT	Likelihood probability based on transition radiation in the TRT. This variable is set to a default value for candidates with $ \eta > 2.01$ due to the limited coverage of the TRT.	eProbabilityHT	DNN
Track-cluster matching	$\Delta\eta$ between the cluster position in the first layer and the extrapolated track	$\Delta\eta_1$	DNN
	$\Delta\phi$ between the cluster position in the second layer of the EM calorimeter and the momentum-rescaled track, extrapolated from the perigee, times the charge q	$\Delta\phi_{res}$	DNN
	Ratio of the cluster energy to the track momentum	E/p	DNN
Kinematics	Transverse energy of the electron measured by the calorimeter system. This variable is not used for discrimination purposes but to give the DNN additional information.	E_T	DNN
	Absolute value of the pseudorapidity of the electron as measured by the calorimeter system. This variable is not used for discrimination purposes but to give the DNN additional information.	$ \eta $	DNN
Reconstruction	Output of an ambiguity resolution algorithm to distinguish objects that are reconstructed as both electrons and photons [30].	Ambiguity type	C