Multivariate analysis of the EW gauge bosons' polarisation at the LHC

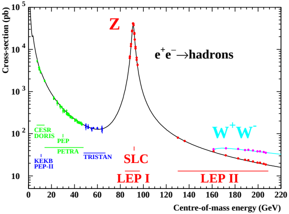
Comité de Suivi Individuel

Mathis DUBAU
Laboratoire d'Annecy de Physique des Particules - ATLAS group

15th July 2024

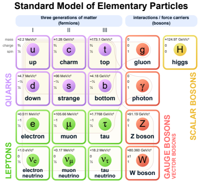
Standard Model

- Theory used in particle physics



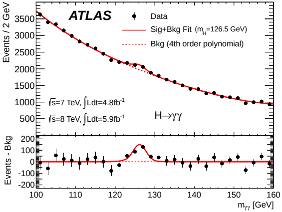
Cross section of $e^+e^- \rightarrow$ hadrons processes as a function of center-of-mass energy [16]

- Successfully confirmed by the discovery of the Higgs Boson in 2012 at the LHC by the ATLAS [14] and CMS [13] experiments



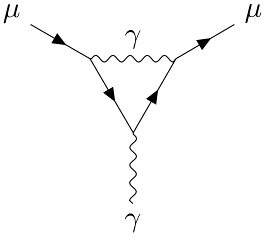
Standard Model

- Highly accurate and extensively tested model

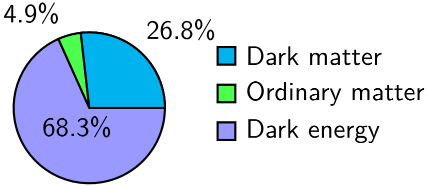


Invariant mass distribution of di-photon candidates for combined data at $\sqrt{s} = 7 : TeV$ and $\sqrt{s} = 8 : TeV$ [14]

- 4.2σ deviation between the Fermilab measurement of the muon's magnetic moment [1] and the prediction by the *Muon g-2 Theory Initiative* [3]



- Lack of a dark matter candidate particle



Energy density distribution of the Universe [9]

Plan

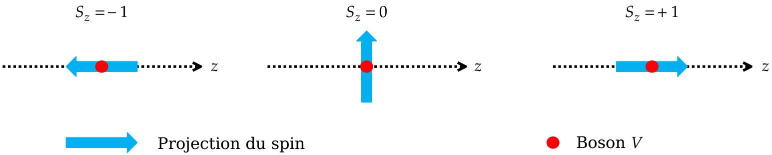
Spin

Intrinsic property of a particle

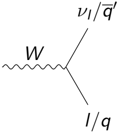
Massive vector bosons

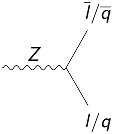
3 degrees of freedom represented by 3 different polarizations, one arise from the higgs mechanism when the bosons acquire mass

$$\epsilon_-^\mu = \frac{1}{\sqrt{2}}(0, 1, -i, 0) \qquad \epsilon_0^\mu = \frac{1}{m_V}(k_z, 0, 0, E) \qquad \epsilon_+^\mu = -\frac{1}{\sqrt{2}}(0, 1, i, 0)$$

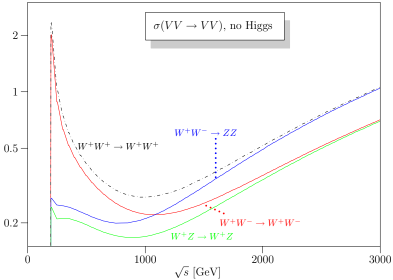


Representation of transverse (T) and longitudinal (0) polarization states

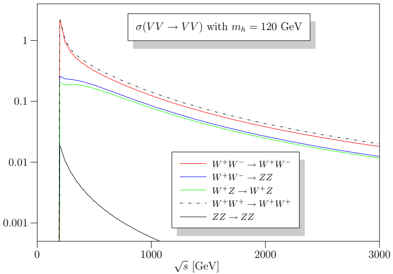




The decay products of W or Z bosons retain traces of the bosons' polarization.



(a) Without Higgs boson



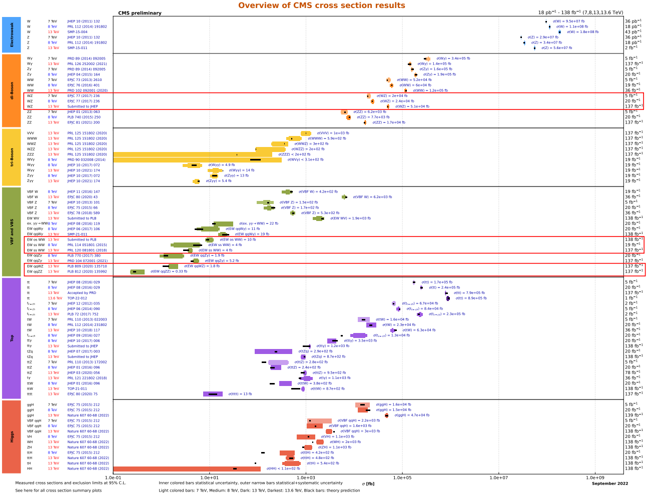
(b) With Higgs boson

Effective cross-section (in *nb*) for five different longitudinally polarized weak interaction gauge boson scattering (VBS *) processes [2]

Measuring polarization in boson scattering (WZ production, for example) provides a direct probe of EW symmetry breaking mechanism.

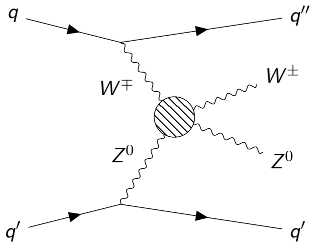
* *Vector Boson Scattering*

Cross Section for *VV* Pairs



Summary of cross section measurements for Standard Model processes by the CMS experiment [18]

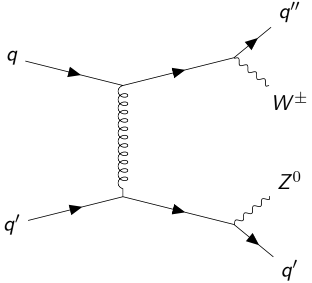
Production of WZ Pairs and 2 Tagging Jets



EW VBS Process - Signal

Decay Channel

$WZjj \rightarrow lll\nu jj$



QCD Process - Background

The leptons considered in this study are electrons and muons

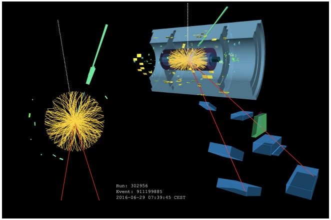
A fist step towards joint-polarisation measurement for WZjj-EW

WZ-EW analysis group

The ATLAS team at LAPP (Emmanuel, Iro, Lucia ...) has been analyzing *WZ* boson pairs for several years, in particular

- the first observation of the production of a *WZ* pair in an electroweak process [6]
- the first observation of the joint polarization states of gauge bosons in the *WZ* production [7]

We also collaborate with the Thessaloniki team on VBS and the Victoria University team on EFT.

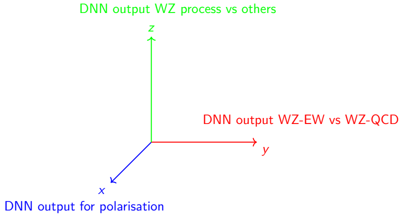


Display of event candidate $WZ \rightarrow e\nu_e\mu\mu$

Steps of the study

- Different machine learning techniques first with simple one with TMVA [20, 10] and then DNN with Tensorflow [11]
- Discrimination between EW process vs QCD process and then polarisation discrimination for TMVA's method, one more discrimination for WZ events versus other process for the DNN (such as tZ, ZZ, $t\bar{t}V$, VVV, ...)

- 2D or 3D map made from the output of the ML techniques and give to a statistic tools to compute significance

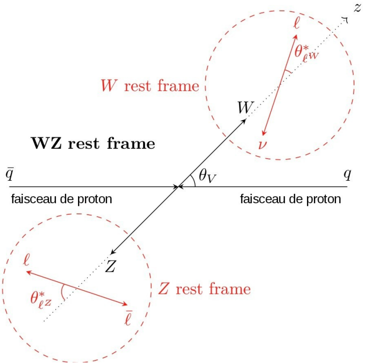


Particle Information

- Leptons
- Bosons
- Tagging Jets

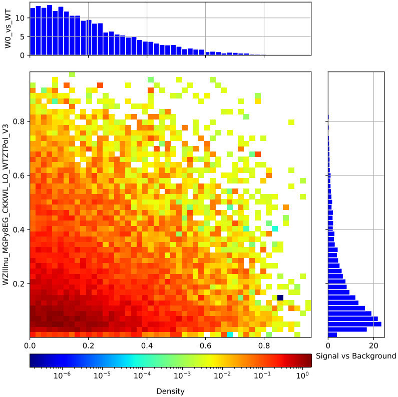
Different Properties

- Kinematics
- Energy
- Centrality
- Number of jets

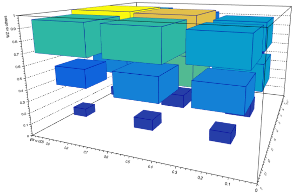


Schematic view of angular observables sensitive to polarization states

Examples of 2D and 3D map



2D map for a $W0Z0$ EW sample



Example of 3D map of the $W0Z0$ sample with the $W0ZX$ vs $WTZX$ DNN

Significance

Significance for the observation of WZjj-EW polarisation process with different methods and variables at $\mathcal{L} = 140\ fb^{-1}$ corresponding at the Run 2 data

Methods	$W0ZX$ vs $WTZX$	$WXZ0$ vs $WXZT$	$W0Z0$ vs others	$WTZT$ vs others	$W0ZT$ vs others	$WTZ0$ vs others
Likelihood	1.39	1.42	0.54	2.76	0.85	0.82
MLP	1.84	1.99	0.67	3.84	1.23	1.32
BDTG	2.04	2.24	0.77	4.08	1.36	1.47
DNN	**2.54**	**2.74**	**0.97**	**4.86**	**1.7**	**1.76**

Significance for the observation of WZjj-EW polarisation process with different methods and variables at $\mathcal{L} = 300\ fb^{-1}$ corresponding at the expected Run 3 data

Methods	$W0ZX$ vs $WTZX$	$WXZ0$ vs $WXZT$	$W0Z0$ vs others	$WTZT$ vs others	$W0ZT$ vs others	$WTZ0$ vs others
Likelihood	1.97	2.01	0.77	3.9	1.21	1.16
MLP	2.6	2.81	0.94	5.42	1.73	1.86
BDTG	2.89	3.17	1.09	5.76	1.93	2.08
DNN	**3.69**	**3.98**	**1.4**	**7.08**	**2.44**	**2.53**

Where X means 0 or T.

Combined Run 2 and Run 3

Significance for the observation of WZjj-EW polarisation process with different methods and variables at $\mathcal{L} = 139 + 300\ fb^{-1}$

Methods	$W0ZX$ vs $WTZX$	$WXZ0$ vs $WXZT$	$W0Z0$ vs others	$WTZT$ vs others	$W0ZT$ vs others	$WTZ0$ vs others
DNN	**4.43**	**4.79**	**1.69**	-	**2.92**	**3.04**

Where X means 0 or T.

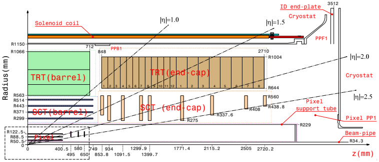
This was done by considering two signal region, one for Run 2 with $\mathcal{L} = 139\ fb^{-1}$ and one for Run 3 with $\mathcal{L} = 300\ fb^{-1}$ who were the same respectivly as the one used on the previous slide for the DNN.

Yet, no control region were used and also the data simulated for Run 3 are exactly the same as the one for Run 2 but rescaled to simulate the increasing luminosity

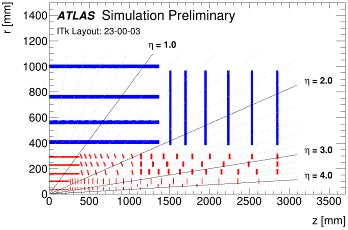
Next step: fractions of polarisation with uncertainty

ATLAS Upgrade Phase 2

Improved $|\eta|$ coverage rate from 2.5 to 4.0 during ATLAS Phase 2 Upgrade, enabling more "forward" electrons/muons to be used.



Current Trajectograph



Inner Tracker (ITk)

Diagram of the quarter ($z > 0, \eta > 0$) of the current Trajectograph and ITk

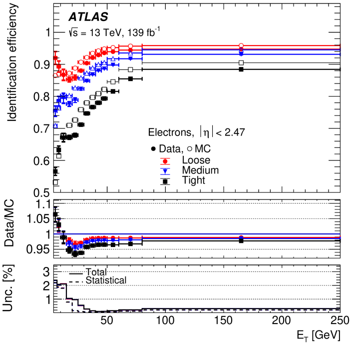
Leptons as final decay with great power of discrimination in the forward region, the ITk upgrade is a welcome one for the VBS analysis we're doing.

Pseudorapidity η

Inclination with respect to the beam, $\eta = 0 \iff$ perpendicular, $\eta = \pm\infty \iff$ parallel

Methodology

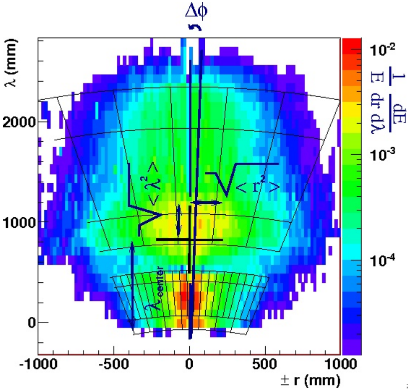
- Trying to set up a p_T-independent identification discriminant in order to make easier the measurement and the treatment of data vs MC discrepancies, as well as the extrapolation to phase space (high p_T) where those can't be measured
- For this, using input variables to a DNN that have little correlation with pT (after employing a decorrelation technique)



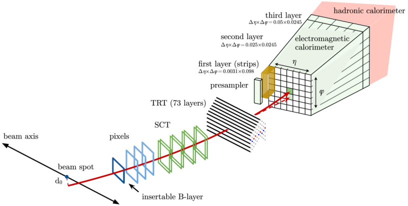
Identification efficiencies of electrons from $Z \longrightarrow ee$ decays as a function of the electron's E_T [5]

- A single DNN can then be trained for all p_T
- To recover and adjust best working points, cuts on this DNN can finally be set in bins of p_T
- Note that in order to adapt to the different detector geometry (boundaries, granularity,...) this process is repeated independently in different bins of $|\eta|$

Variables



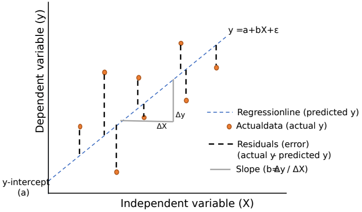
Clusters moments varaibles



Current Trajectograph

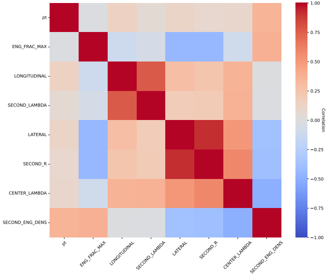
Example of variables we can get from the detectors

p_T decorrelation

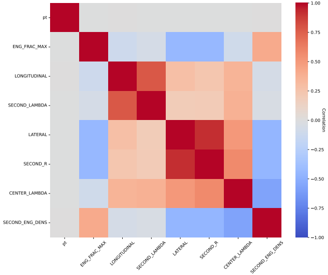


Linear regression and residuals

decorrelation between reconstructed p_T and the seven clusters moments by training a linear regression model to fit p_T versus the 7 C.M (for signal only). Then we take the residuals to subtract them to the C.M (background and signal). The p_T variable is then unchanged, and is uncorrelated from the 7 C.M



(a) Before uncorrelation

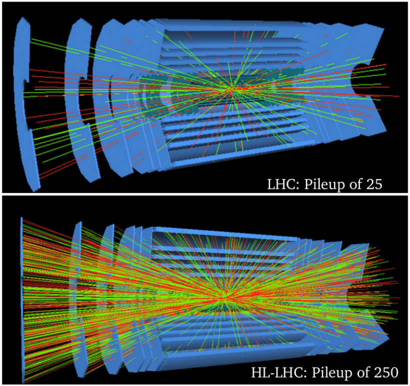


(b) After uncorrelation

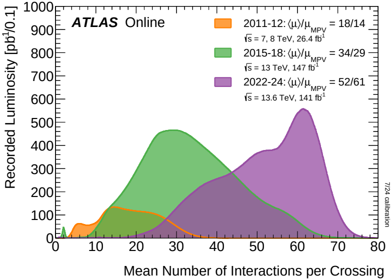
Correlation matrix between p_T and the seven clusters moments for the signal before and after uncorrelated them with a linear regression model example for $2.7 \leq |\eta| \leq 3.2$

As expected, the decorrelation process allows for using in the DNN these 7 input variables independent of pT

Pile-up



Representation of the pile-up effect

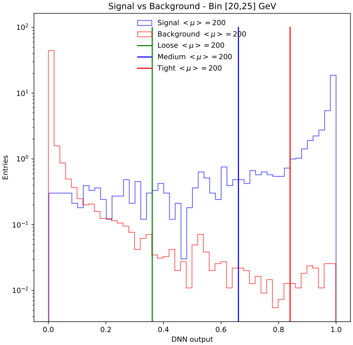


Luminosity-weighted distribution of the mean number of interactions per crossing for *pp* collisions for Run 1, 2 and 3

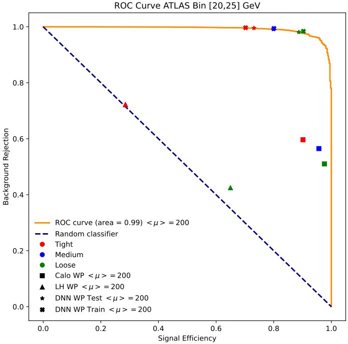
Pile-up $<\mu>$

Number of proton-proton collisions per bunch crossing at the interaction point

Bin [20;25] for data with $< \mu > \simeq 200$ - $2.7 \leq |\eta| \leq 3.2$

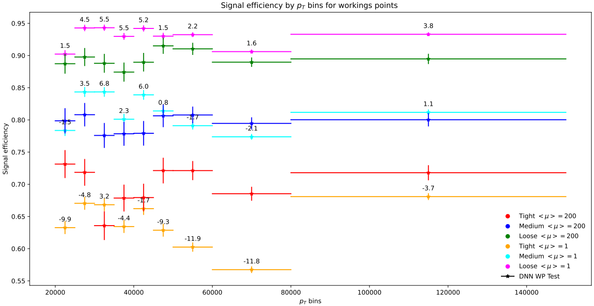


DNN prediction for signal and background with cut on signal efficiency



ROC curve and various workings points

DNN WP Signal efficiency - $2.7 \leq |\eta| \leq 3.2$



Signal efficiency for DNN WP for data test at $<\mu>= 200$ and $<\mu>= 1$

The number above the DNN WP at $<\mu>= 1$ is the difference in percentage between the one at $<\mu>= 200$
On average, signal efficiency at $<\mu>= 1$ is a few percent higher than $<\mu>= 200$ for loose and medium
The sensitivity of the signal selection to pile-up is only a few percent

Formations

- Research ethics - 15h
- Fundamentals of Big Data - 24h
- Gif School 2023 - 24h
- PhD and career development - 24h
- Introduction to parallel computing - 36h
- European School of High Energy Physics - end of 2024

Work on the side

- 21h30 of lectures on Introduction to python for 1st year student at the USMB
- Vulgarisation scientifique pour Fête de la science of LAPP 2023 and at Mercredi du LAPP
- CERN Guide
- Shift in the control room for the Calorimeter / Forward detector desk - 208h
- Develop a website to visualise our datasets
- E/gamma Workshop in Valencia in the context of my QT

The QT

- Completing the QT, we will request more statistic to consolidate the first result presented here. There will be an implementation of the DNN inside the Athena framework

The analysis

- Continuing the VBS study by increasing the MC statistic
- Use fraction of polarisation instead of significance
- Definition of control regions and study of associated systematic uncertainties through the statistical treatment

Work on the side

- Continue to dispense python lectures for 1st year student at USMB

References

[1] B. Abi et al. “Measurement of the Positive Muon Anomalous Magnetic Moment to 0.46 ppm”. In: *Phys. Rev. Lett.* 126.14 (2021), p. 141801. DOI: 10.1103/PhysRevLett.126.141801. arXiv: 2104.03281 [hep-ex].

[2] Ana Alboteanu, Wolfgang Kilian, and Jürgen Reuter. “Resonances and unitarity in weak boson scattering at the LHC”. In: *Journal of High Energy Physics* 2008.11 (Oct. 2008), pp. 010–010. DOI: 10.1088/1126-6708/2008/11/010.

[3] T. Aoyama et al. “The anomalous magnetic moment of the muon in the Standard Model”. In: *Phys. Rept.* 887 (2020), pp. 1–166. DOI: 10.1016/j.physrep.2020.07.006. arXiv: 2006.04822 [hep-ph].

[4] François Chollet et al. *Keras*. https://keras.io. 2015.

[5] ATLAS Collaboration. “Electron and photon efficiencies in LHC Run 2 with the ATLAS experiment”. In: *Journal of High Energy Physics* 2024.5 (May 2024). ISSN: 1029-8479. DOI: 10.1007/jhep05(2024)162. URL: http://dx.doi.org/10.1007/JHEP05(2024)162.

[6] ATLAS Collaboration. “Observation of electroweak $W^{\pm}Z$ boson pair production in association with two jets in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector”. In: *Physics Letters B* 793 (June 2019), pp. 469–492. DOI: 10.1016/j.physletb.2019.05.012.

[7] ATLAS Collaboration. *Observation of gauge boson joint-polarisation states in $W^{\pm}Z$ production from pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*. 2022. arXiv: 2211.09435 [hep-ex].

[8] Gessinger, Paul et al. “The Acts project: track reconstruction software for HL-LHC and beyond”. In: *EPJ Web Conf.* 245 (2020), p. 10003. DOI: 10.1051/epjconf/202024510003. URL: https://doi.org/10.1051/epjconf/202024510003.

[9] G. Hinshaw et al. “Nine-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Cosmological Parameter Results”. In: *Astrophys. J. Suppl.* 208 (2013), p. 19. DOI: 10.1088/0067-0049/208/2/19. arXiv: 1212.5226 [astro-ph.CO].

[10] Andreas Hoecker et al. “TMVA: Toolkit for Multivariate Data Analysis”. In: *PoS* ACAT (2007), p. 040. arXiv: physics/0703039.

[11] Martín Abadi et al. *TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems*. Software available from tensorflow.org. 2015. URL: https://www.tensorflow.org/.

[12] Sascha Mehlhase. “ATLAS detector slice (and particle visualisations)”. In: (2021). URL: https://cds.cern.ch/record/2770815.

[13] “Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC”. In: *Physics Letters B* 716.1 (Sept. 2012), pp. 30–61. DOI: 10.1016/j.physletb.2012.08.021. URL: https://doi.org/10.1016/j.physletb.2012.08.021.

[14] “Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC”. In: *Physics Letters B* 716.1 (Sept. 2012), pp. 1–29. DOI: 10.1016/j.physletb.2012.08.020. URL: https://doi.org/10.1016/j.physletb.2012.08.020.

[15] Louis Portales. “Observation of electroweak WZjj production and studies on pile-up mitigation with the ATLAS detector”. Theses. Université Savoie Mont Blanc, Oct. 2020. URL: https://theses.hal.science/tel-03550156.

[16] “Precision electroweak measurements on the Z resonance”. In: *Phys. Rept.* 427 (2006), pp. 257–454. DOI: 10.1016/j.physrep.2005.12.006. arXiv: hep-ex/0509008.

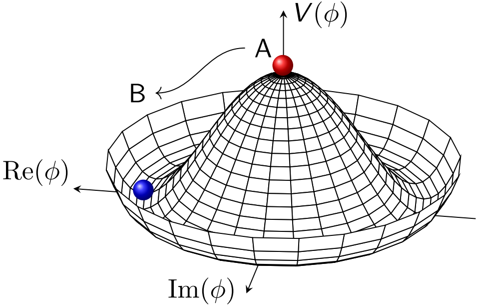
[17] Luka Selem. “Measurement of gauge boson joint-polarisation states in WZ pair production at the LHC with the ATLAS detector. Mesure des états de co-polarisation des bosons de jauge en production de paire WZ au LHC avec le détecteur ATLAS”. Presented 27 Sep 2022. Savoie U., 2022. URL: http://cds.cern.ch/record/2841405.

[18] *Summaries of CMS cross section measurements*. https://twiki.cern.ch/twiki/bin/view/CMSPublic/PhysicsResultsCombined. June 2023.

[19] “The ATLAS Experiment at the CERN Large Hadron Collider”. In: *JINST* 3 (2008). Also published by CERN Geneva in 2010, S08003. DOI: 10.1088/1748-0221/3/08/S08003. URL: https://cds.cern.ch/record/1129811.

[20] Jan Therhaag. “TMVA Toolkit for multivariate data analysis in ROOT”. In: *PoS* ICHEP2010 (2010). Ed. by Bernard Pire et al., p. 510. DOI: 10.22323/1.120.0510.

[21] *TRExFitter*. https://gitlab.cern.ch/TRExStats/TRExFitter/. 2023.



Representation of the Higgs potential
$V(\phi) = \mu^2 \phi^* \phi + \lambda (\phi^* \phi)^2$

Spontaneous Symmetry Breaking

$$\langle \phi \rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v \end{pmatrix}$$

$$\Downarrow$$

Quantum Fluctuations

$$\phi(x) = \frac{1}{\sqrt{2}} \begin{pmatrix} \phi_1(x) + i\phi_2(x) \\ v + \phi_4(x) + i\phi_3(x) \end{pmatrix}$$

$$\Downarrow$$

Unitary Gauge

$$\phi(x) = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + h(x) \end{pmatrix}$$

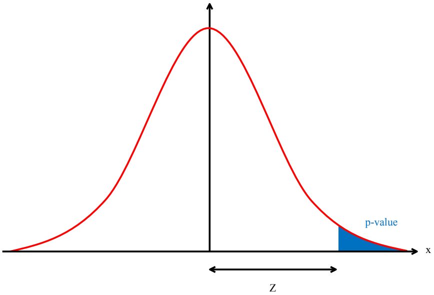
The gauge bosons acquire mass and absorb a Goldstone boson, which manifests as a third polarization state.

Significance and *p-value*

p-value

Probability of obtaining a value t in the region of compatibility with H_0 as extreme or more extreme than the value observed in the real data.

$$p = \int_{t_{obs}}^{\infty} f(t|H_0)dt$$



Relation between significance Z and the *p-value*

Discovery in HEP

$Z = 5$ or $5\sigma \iff$ *p-value* $= 2.87 \times 10^{-7}$

Phase space

Phase space definition

Variables	Fiducial WZjj-EW
Lepton $\lvert\eta\rvert$	< 2.5
p_T of l_Z, p_T of l_W [GeV]	> 15, > 20
m_Z range [GeV]	$\lvert m_Z - m_Z^{PDG}\rvert < 10$
m_T^W [GeV]	≥ 30
$\Delta R(l_Z^-, l_Z+)$, $\Delta R(l_Z, l_W)$	> 0.2, > 0.3
p_T two leading jets [GeV]	> 40
$\lvert\eta_j\rvert$ two leading jets	< 4.5
Jets multiplicity	≥ 2
$\eta_{j1} \cdot \eta_{j1}$	< 0
m_{jj} [GeV]	> 500
$\Delta R(j, l)$	> 0.3
$N_{b-quark}$	$= 0$

Tables of variables used

Usage	Symbol	Description	Type information
Variables of the DNN of differentiation of polarizations VBS EW [17]	$\Delta\phi(l^W, l^\nu)$	Difference in azimuthal angle ϕ of leptons from W	Information about leptons
	$\Delta\phi(l_1^Z, l_2^Z)$	Difference of azimuthal angle ϕ of leptons from Z	
	$p_T^{W_l, Z_l}$	Transverse impulse of leptons from W and Z	
	E_T^{miss}	Missing transverse energy	
	p_T^{WZ}	Transverse impulse of the WZ system	Information on jets and bosons
Variables of the BDTG of differentiation EW processes vs QCD [15]	$\Delta R(j_1, Z)$	Angular separation between the first tagging jet and the Z boson	
	$R_{p_T}^{\text{hard}}$	Transverse component of the vector sum of the moments of the hard objects in the final state of events (leptons and jets), divided by the sum of their transverse moments.	
	ζ_{lep}	Centrality of leptons out of respect for the di-jet system	
	ζ_{jet}	Centrality of the jets	
	m_{jj}	Mass of the di-jet system	Information on the tagging jets
	N_{jets}	Multiplicity of jets	
	$p_T^{j_1, j_2}$	Transverse impulse of the two tagging jets	
	η_{j_1}	Pseudorapidity of jet 1	
	$\Delta\eta_{jj}$	Difference in pseudorapidity between the two tagging jets	
	$\Delta\phi_{jj}$	Difference in azimuth angle between the two tagging jets	
	η_W	Boson pseudorapidity W	Boson information
	m_T^{WZ}	Transverse mass of system WZ	
	$p_T^{W,Z}$	Transverse impulse of the W and Z bosons	
Variables common to DNN and BDTG	$\lvert y_{l^W} - y_Z \rvert$	Difference in velocity between the Z boson and the lepton from the W boson	
Additional variables	$\cos\theta_{W,Z,V}$	cosine of the angle between the emission of the two leptons in the frame of reference of W, Z and in the frame of reference at rest of WZ	
	r^{21}	p_T of the leading Boson p_T divided by p_T of the sub-leading Boson p_T	
	η_{W_l, Z_l}	Pseudorapidity of the three leptons from W and Z	Information on the leptons
	ϕ_{W_l, Z_l}	Azimuthal angle ϕ of the three leptons from the W and the Z	
	$\cos\theta_{Vjj}$	cosine of the angle between the two tagging jets in the WZ frame of reference at rest	Information on the tagging jets
	η_{j_2}	Pseudorapidity of jet 2	
	ϕ_{j_1, j_2}	Azimuthal angle of the two tagging jets	
	E_{j_1, j_2}	Energy of the two tagging jets	
	$\langle p_T \rangle$	Average p_T of the 2 tagging jets	
	$A_{p_T^{jj}}$	Asymmetry between the pT of the 2 tagging jets	

DNN

We use **`Tensorflow`** [11] and his API **`Keras`** [4] to compute deep neural networks
The dataset, is normalised for better performance with following normalization:

$$x_{norm} = \frac{x - x_{max}}{x_{max} - x_{min}}$$

to scale them between 0 and 1
We labeled signal as 1 and background as 0 as so use the binary cross entropy function:

$$\mathscr{L} = \frac{-1}{N} \sum_{i=1}^{N} x_i \times \ln \hat{x}_i + (1 - x_i) \times \ln (1 - \hat{x}_i)$$

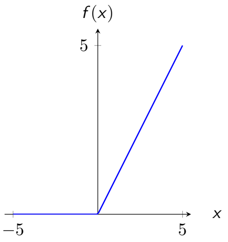
We split the dataset as 80% of it for training and 20% for validation

We then search for the best hyperparemeters with a Bayesian optimizer from `KerasTunner` for the following values:

- number of hidden layers $\in [1 - 15]$
- number of neuron in hidden layers $\in [32 - 256]$
- learning rate $\in [10^{-2}, 10^{-3}, 10^{-4}]$

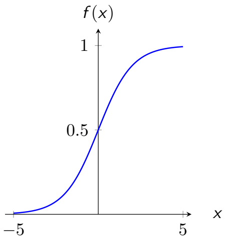
The optimizer for the loss function is `Adam`. The input layer is made of 41 neurons and the output one of only 1. Each layer except the output one has a ReLu function as activation function, the output one has a sigmoid function.

$$ReLU(x) = \begin{cases} x & \text{if } x > 0 \\ 0 & \text{else} \end{cases}$$



ReLU (*Rectified Linear Unit*) function

$$\sigma(x) = \frac{1}{1 + \exp^{-x}}$$



Sigmoid function

The `KerasTunner` will then compute 100 different model and test them on 10 epochs (with an early stopper focused on loss value with a patience of 3) and watch the best values.

The 100 steps were not done for the following results, only 10 to give a rough idea

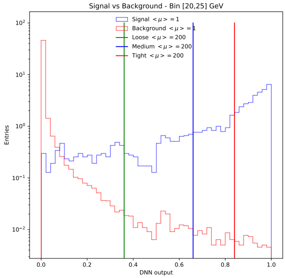
Variables

Variables names	Variables information
time	HGTD time information
calo_eta	η from CaloCluster
calo_phi	ϕ from CaloCluster
calo_E_2	Deposit energy on the 2nd layer of the EMCalo
calo_E_3	Deposit energy on the 3rd layer of the EMCalo
track_eta	η from trackParticle
track_phi	ϕ from trackParticle
pixels	Number of pixels hits
strips	Number of strips hits
ENG_FRAC_MAX	f_{em} Fraction of the cluster energy that is deposited in the most energetic cell of the cluster
LONGITUDINAL	Shower shape in the clusters' longitudinal direction, based on the distance of each cell to the shower core
SECOND_LAMBDA	$< \lambda^2 >$ Second moment in λ - the distance of each cell to the cluster center along the shower axis
LATERAL	Lateral moment of the shower taking into account the two most energetic cells (which constitutes the shower core)
SECOND_R	$< r^2 >$ Second moment in r - the radial distance of each cell to the shower axis
CENTER_LAMBDA	λ_{center} the distance of the shower center from the front of the calorimeter along the shower axis
SECOND_ENG_DENS	$< \rho^2 >$ Second moment in energy density
delta_eta2	Delta squared in η between caloCluster and trackParticle
delta_phi2	Delta squared in ϕ between caloCluster and trackParticle
delta_phi_rescaled2	
delta_phi_last	

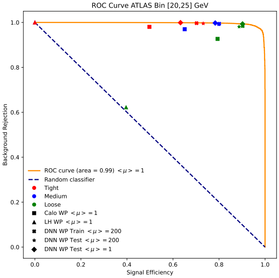
Variables used for training

The six first Cluster Moments (f_{em}, `longitudinal`, $< \lambda^2 >$, `lateral`, $< r^2 >$ and λ_{center}) are the ones to ID the electron in the forward region presently

Bin [20;25] for data with $<\mu>= 1$ - $2.7 \leq |\eta| \leq 3.2$



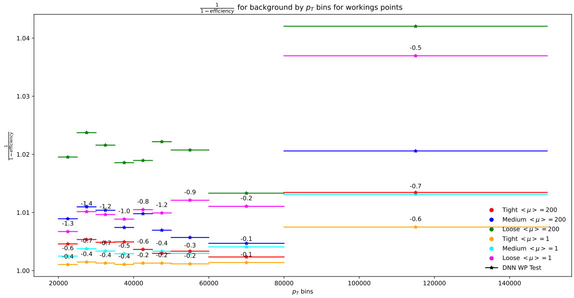
DNN prediction for signal and background with working points defined for $<\mu>= 200$



ROC curve and various workings points

Here testing the $<\mu>= 200$ WP on a $<\mu>= 1$ sample $\longrightarrow$ good stability with only a few percents change of efficiencies

DNN WP Background efficiency - $2.7 \leq |\eta| \leq 3.2$



Inverse background efficiency for DNN WP for data test at $<\mu>=200$ and $<\mu>=1$

The number above the DNN WP at $<\mu>=1$ is the difference in percentage between the one at $<\mu>=200$

On average, background efficiency at $<\mu>=1$ is the same than $<\mu>=200$

The sensitivity of the background selection to pile-up is the same