Polarised boson tagging at the LHC with machine-learning methods

Polarized Perspectives:
Tagging and Learning in the SM

Vienna 21st February 2025

Michele Grossi, PhD

CERN QTI

CERN

QUANTUM
TECHNOLOGY
INITIATIVE

Amplitude-assisted tagging of longitudinally polarised bosons using wide neural networks

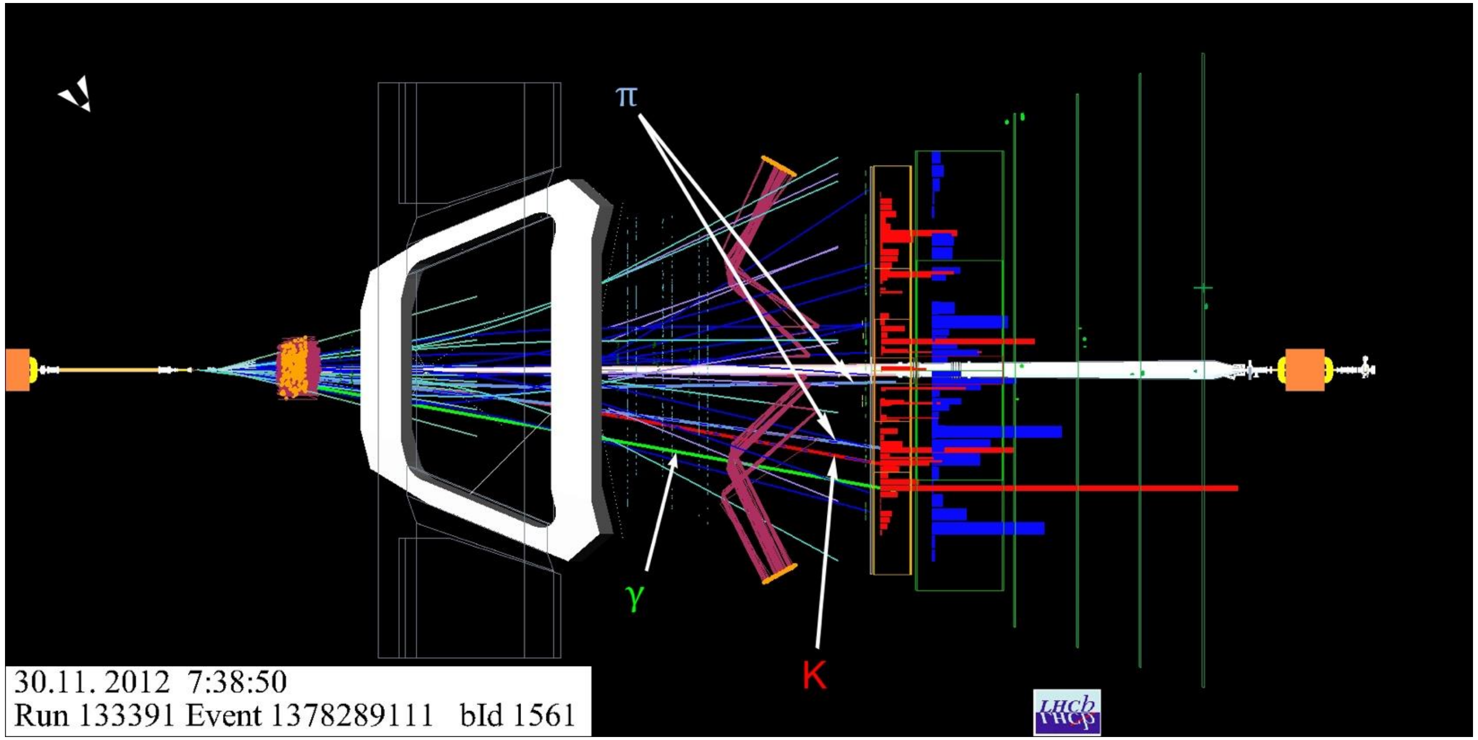
Paper:

Grossi, Incudini, Pellen, Pellicciolli - **Eur. Phys. J. C (2023) 83:759**

Slides: revisited material from my co-authors Mathieu Pellen and Giovanni Pellicciolli
THANKS

- **Problem Definition**
- **ML & Polarization Tagging**
- **Results**
- **Summary & Outlook**

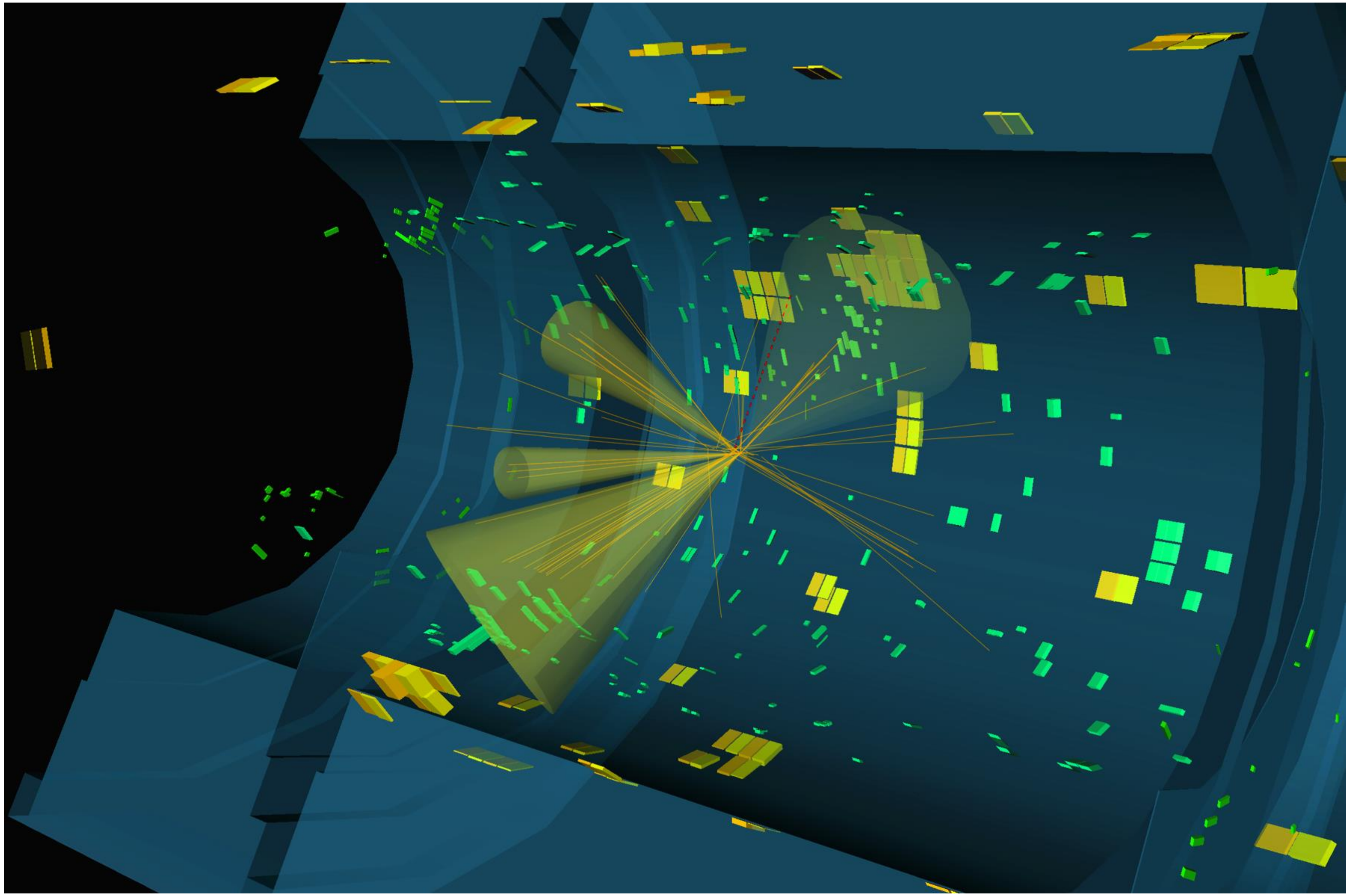
- **Problem Definition**
- **ML & Polarization Tagging**
- **Results**
- **Summary & Outlook**



Can we design an original method to extract polarisation?

Maybe...

Can we design an original method to extract polarisation fractions using the maximal information encoded in the **amplitude**?



YES

Polarization of weak boson

- Important probes of SM gauge and (extended) Higgs sectors
- Polarisation of gauge bosons related to Electroweak symmetry breaking

- The cross-section and angular distribution of longitudinal polarizations are particularly sensitive to beyond standard model (BSM) physics

- Special interest in di-boson (inclusive, VBS, Higgs decays)

The theoretical study and the experimental extraction of such pseudo-observables is thus of prime importance for the present and upcoming physics programme of the LHC

Math – Polarization of weak boson

$$\mathcal{M} = \overset{\text{PROD}}{\mathcal{M}^{\mathcal{P}}_{\mu}} \frac{i}{k^2 - M_V^2 + i\Gamma_V M_V} \underbrace{\left(-g^{\mu\nu} + \frac{k^\mu k^\nu}{M_V^2}\right)}_{\to \sum_{\lambda=1}^{4} \varepsilon^{\mu}_{\lambda}(k)\varepsilon^{\nu *}_{\lambda}(k)} \overset{\text{DEC}}{\mathcal{M}^{\mathcal{D}}_{\nu}}$$

generic (unpolarised) amplitude featuring a resonant gauge boson decaying into a lepton-neutrino pair

$$\sum_{\lambda=1}^{4} \varepsilon^{\mu}_{\lambda}(k)\varepsilon^{\nu *}_{\lambda}(k)$$

polarisation vectors of the massive gauge boson (not Lorentz invariant: defined in a specific frame)

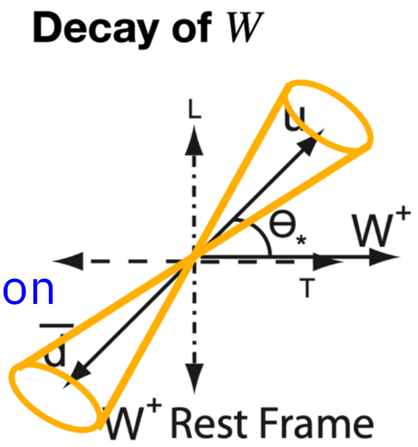
...the unpolarised cross section (measured experimentally)

$$\Rightarrow \quad |\mathcal{M}|^2 = \underbrace{\boxed{|\mathcal{M}_{\mathrm{L}}|^2} + |\mathcal{M}_{\mathrm{T}}|^2}_{\text{Pol. AMP.}} + \underbrace{2\,\mathrm{Re}\,(\mathcal{M}^*_{\mathrm{L}}\mathcal{M}_{+}) + 2\,\mathrm{Re}\,(\mathcal{M}^*_{\mathrm{L}}\mathcal{M}_{-})}_{\text{interf.}}$$

Polarisation state of an unstable particle is not directly accessible in the detectors and the information about it can only be reconstructed (in a probabilistic way) from the stable decay products —> pseudo-observable ('quantum') angular-coefficient extraction

Polarization of weak boson

We cannot directly measure polarisations of EW bosons. But…



Decay-product angular distributions reflect polarisation state of the decayed V boson

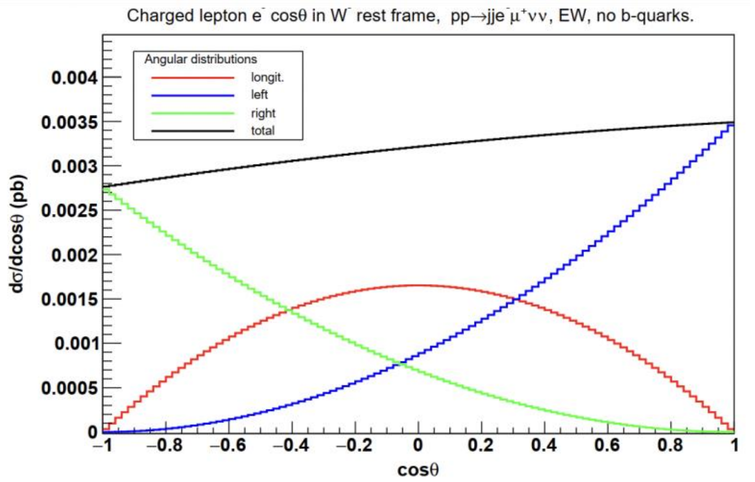
Ex: differential cross-section with no lepton cuts for W->lv is:

$$\frac{1}{\sigma}\frac{d\sigma}{d\cos\theta}(W^\pm \to l^\pm \nu) = \frac{3}{4}f_0 \sin^2\theta + \frac{3}{8}f_R(1 \pm \cos\theta)^2 + \frac{3}{8}f_L(1 \mp \cos\theta)^2$$

where are W polarisation fractions

θ *is the lepton polar angle in the W rest frame wrt W direction in the lab frame*

Cross section (polarization fractions) can be written in terms of *Legendre Polynomials* of decay products



[A. Ballestrero, E. Maina, and G. Pelliccioli; W boson polarization in vector boson scattering at the LHC; JHEP 03 (2018)]

→ Boson polarisation can be measured as angular distributions of particles produced in the decay process from unpolarized VBS process

Grossi et al. 10.1140/epjc/s10052-020-08713-1

- Problem Definition
- ML & Polarization Tagging
- Results
- Summary & Outlook

Machine Learning – not a newcomer

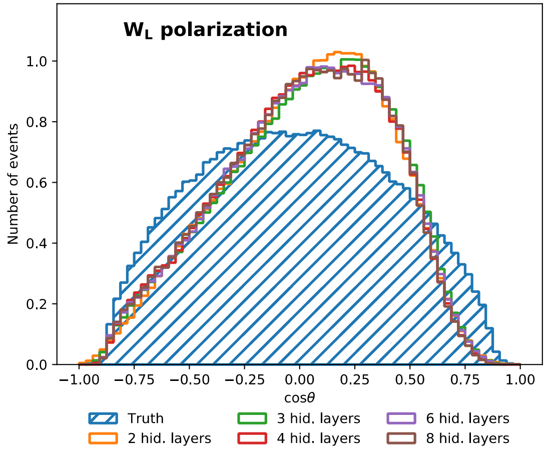
We cannot directly measure polarisations of EW bosons. But…

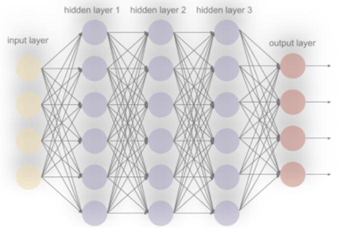
- Various neural-network strategies used for polarisation extraction from leptonic decays in VBS events (lower BR, neutrino reco) [Searcy et al. 1510.01691, Lee et al. 1812.07591, 1908.05196, Grossi et al. 2008.05316, Li et al. 2010.13281, 2109.09924]
- Machine-learning (ML) approaches also used to extract polarisations from hadronic decays (extract correct boson jet, QCD effect)[Kim Martin 2102.05124].

Mostly use kinematical features to reconstruct polarisation-sensitive angles or directly extract polarisation fractions with DNN score.

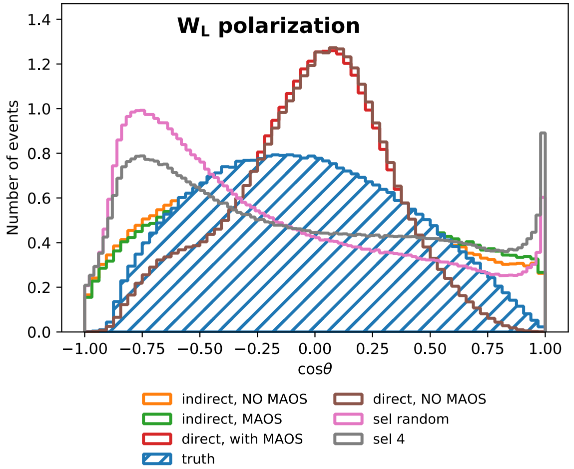
The VBS case… $\cos(\theta)$

- Semi-leptonic: $pp \to j\mu^+\nu_\mu$





- Full-leptonic: $pp \to jje^+\nu_e\mu^+\nu_\mu$



Grossi et al. 10.1140/epjc/s10052-020-08713-1

Machine Learning – not a newcomer

- We can perform fits of LHC data with polarised templates

  → extraction of parameters based on **theory input**

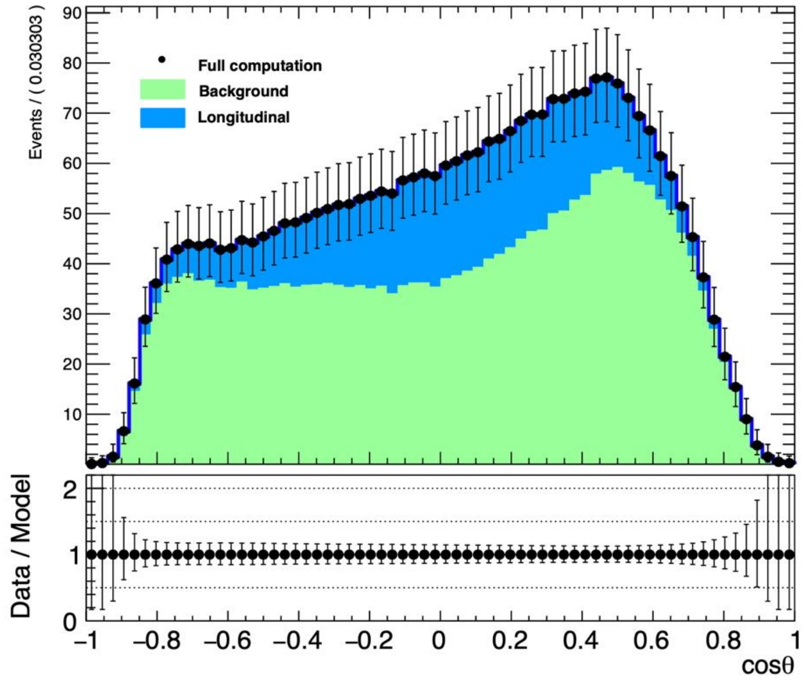
Open questions

- Should the two signatures be fitted separately or together?
- How should theoretical uncertainties be taken into account in the fit?
- How should one define the overall uncertainty on the fit of the polarisation fractions?
- What are the optimal observables?

→ Answers in collaboration with experimental collaborations

Drawbacks

- Restricted to given observables
- Interpretation at the integrated level



Grossi et al. 10.1140/epjc/s10052-020-08713-1

Polarization tagging

New idea for polarisation extraction:
‘Amplitude-assisted tagging of longitudinally polarised bosons using wide neural networks’

use AMPLITUDE to extract parameters for the theory

✓

- no restricted to given observables (no obs)
- event-by-event interpretation (fully differential)
- can be applied to any physics problem that can be cast in one single, bounded, ratio (next)

Not to be confused with
(Similar/connected ideas):

- matrix-element method
  [Kondo; J. Phys. Soc. Jap. 57 (1988) 4126-4140 / 60 (1991) 836–844.]
- optimal-observable method
  [Diehl, Nachtmann; Z. Phys. C 62 (1994) 397-412, hep-ph/9603207], [Janot; 1503.01325]
- MELA (Matrix Element Likelihood Approach)
  [Gao, Gritsan, Melkikov, Schulze, et al.; 1001.3396, 1208.4018,1309.4819, 1606.03107]

Polarization tagging

How to tag an unpolarised event as longitudinal.

$$f_L(\mathcal{O}) = \frac{d\sigma_L}{d\mathcal{O}} \Big/ \frac{d\sigma_{unp}}{d\mathcal{O}}$$

At the event-by-event/phase-space–point level, at leading order is equivalent to →

longitudinal

$$r_L = \frac{|M_L|^2}{|M|^2}$$

unpolarized

r_L would be the optimal observable

BUT, we don't measure it

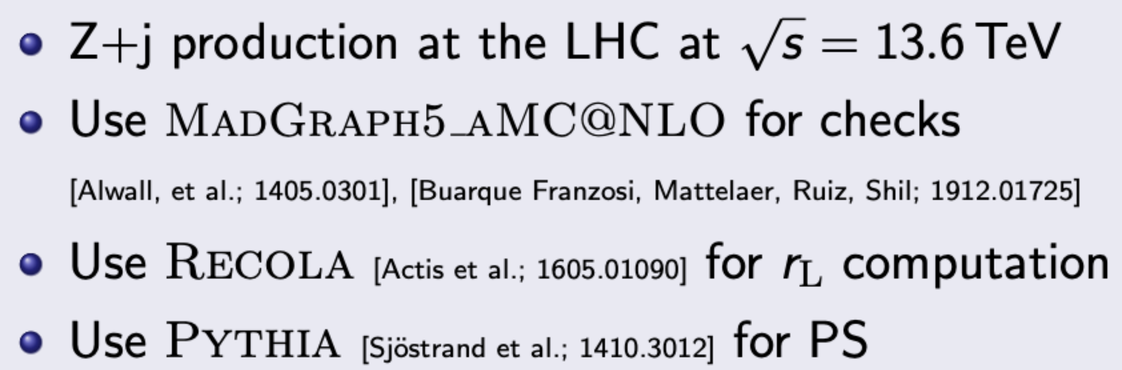
TODO: Unpolarised events → reweight/sample according to truth r_L → longitudinal events (*the probability for an event to be longitudinally polarised*)

Polarization tagging - using wide neural networks

- r_L requires knowledge of all momenta (initial and final), partonic process and the PDF (+ PS)
- experimentally we have access to partial information about the unpolarised process (final-state only, no flavour structure).

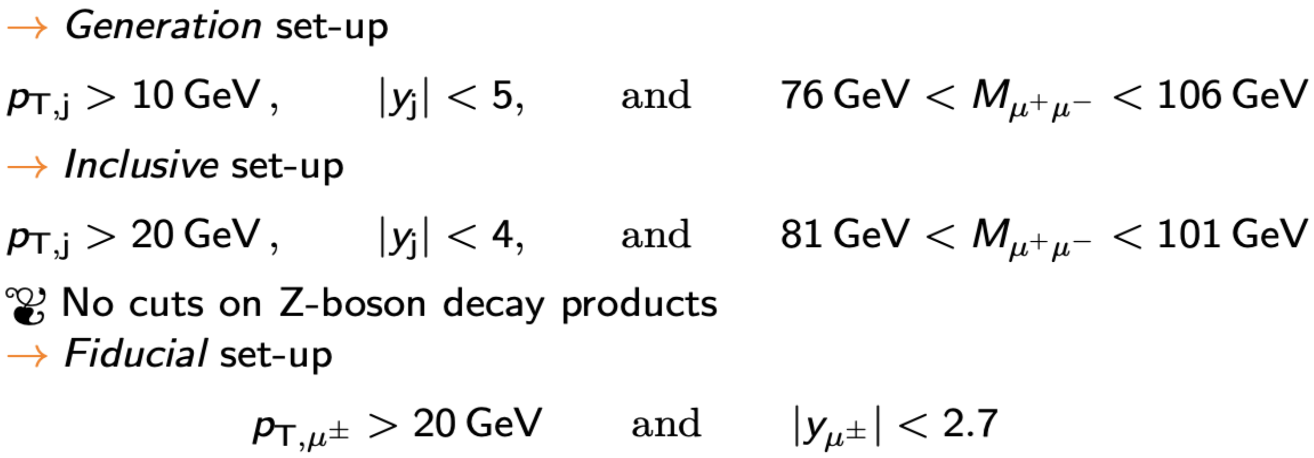
→ bypass this lack of information by using a neural network (NN)

Polarization tagging – setting



- Z+j production at the LHC at $\sqrt{s} = 13.6$ TeV
- Use MadGraph5_aMC@NLO for checks

  [Alwall, et al.; 1405.0301], [Buarque Franzosi, Mattelaer, Ruiz, Shil; 1912.01725]
- Use Recola [Actis et al.; 1605.01090] for r_{L} computation
- Use Pythia [Sjöstrand et al.; 1410.3012] for PS



→ *Generation* set-up

$$p_{\mathrm{T,j}} > 10\,\mathrm{GeV}\,, \qquad |y_{\mathrm{j}}| < 5, \qquad \text{and} \qquad 76\,\mathrm{GeV} < M_{\mu^+\mu^-} < 106\,\mathrm{GeV}$$

→ *Inclusive* set-up

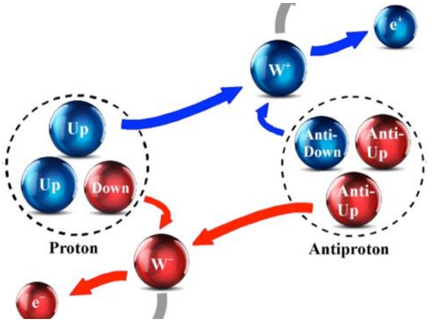
$$p_{\mathrm{T,j}} > 20\,\mathrm{GeV}\,, \qquad |y_{\mathrm{j}}| < 4, \qquad \text{and} \qquad 81\,\mathrm{GeV} < M_{\mu^+\mu^-} < 101\,\mathrm{GeV}$$

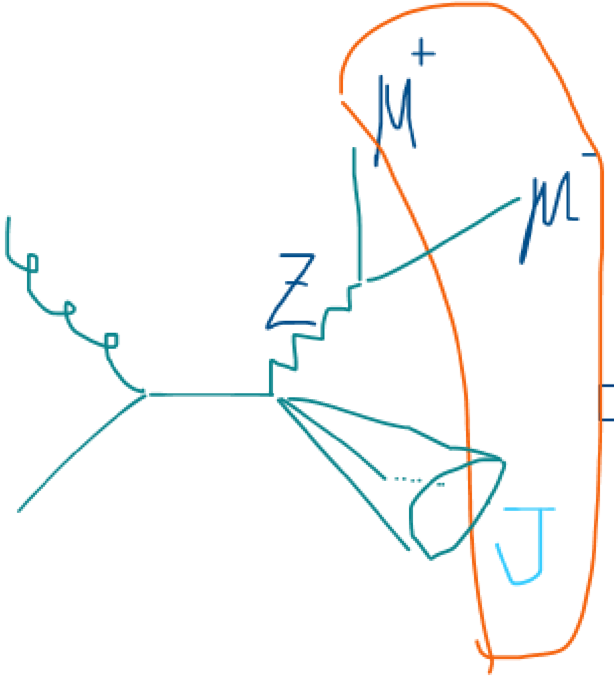
No cuts on Z-boson decay products

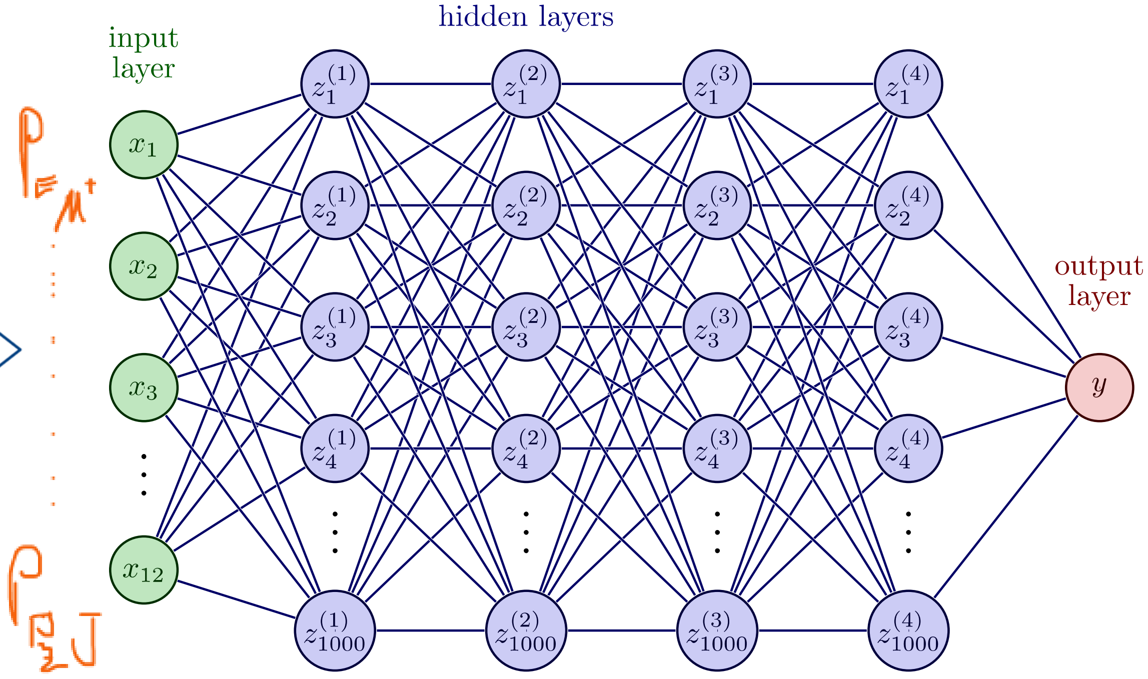
→ *Fiducial* set-up

$$p_{\mathrm{T},\mu^\pm} > 20\,\mathrm{GeV} \qquad \text{and} \qquad |y_{\mu^\pm}| < 2.7$$

Polarization tagging - using wide neural networks



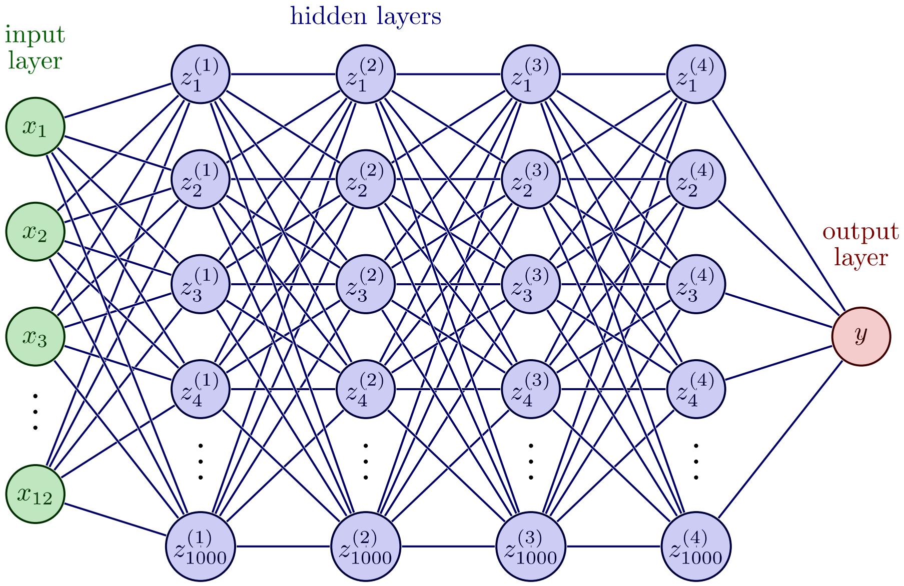




$p_{\mu^+}, p_{\mu^-}, p_j;\ r_L$

$p_{\mu^+}, p_{\mu^-}, p_j;\ \tilde{r}_L$

Polarization tagging - using wide neural networks



PyTorch library and RMSprop algorithm for gradient-descent optimisation

Physics DETAILS

- NN for fixed-order events
- NN for parton-showered events
- training in inclusive setup, testing in inclusive and fiducial one.

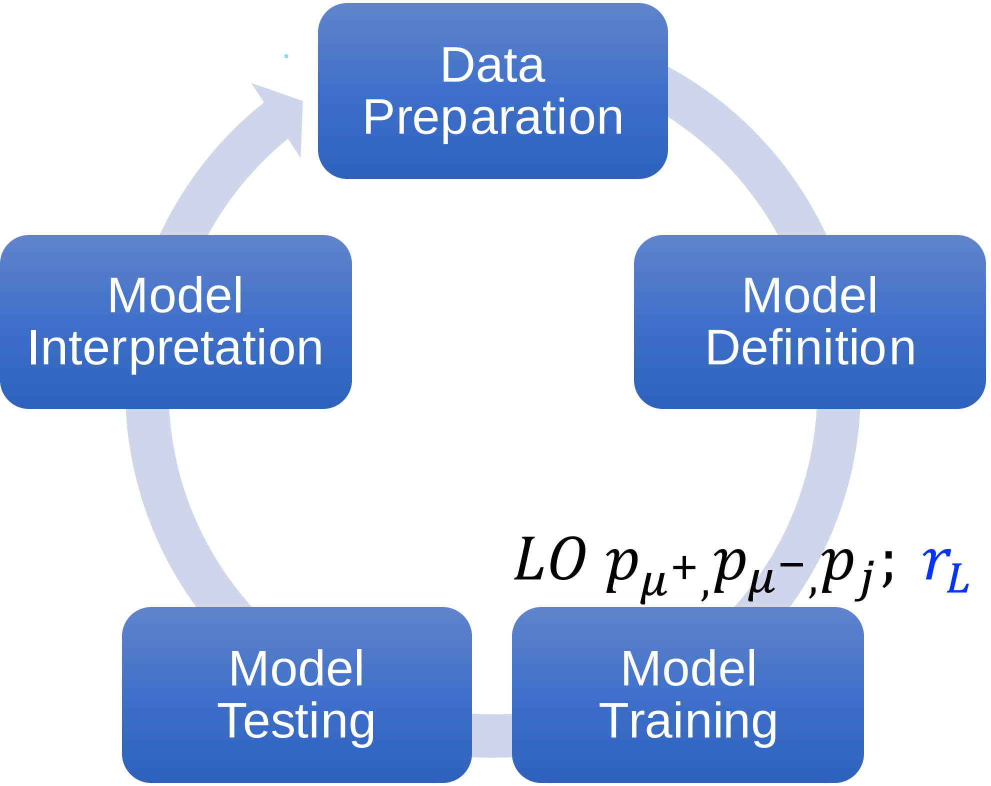
NN DETAILS

- Optimal-model width-to-depth ratio found to be 1000 : 5 → the NN is rather wide.

shallow and wide are easier to train (no caothic dynamics), but potential lost of non-linear expressivity (i.e. linear model).

- Problem Definition
- ML & Polarization Tagging
- Results
- Summary & Outlook

Polarization tagging - LO


Data Preparation

Model Definition

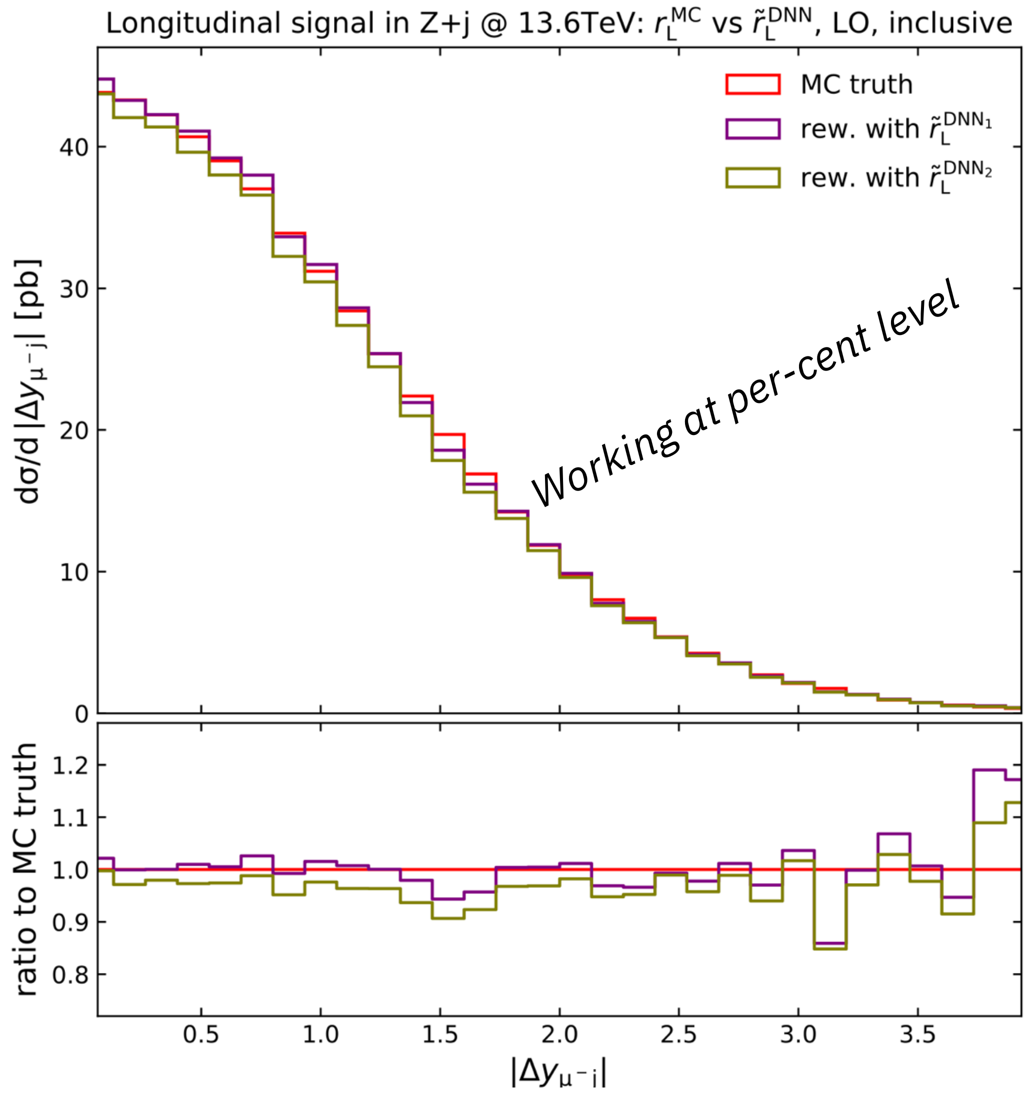
Model Training

Model Testing

Model Interpretation

$LO\ p_{\mu^+}, p_{\mu^-}, p_j;\ r_L$

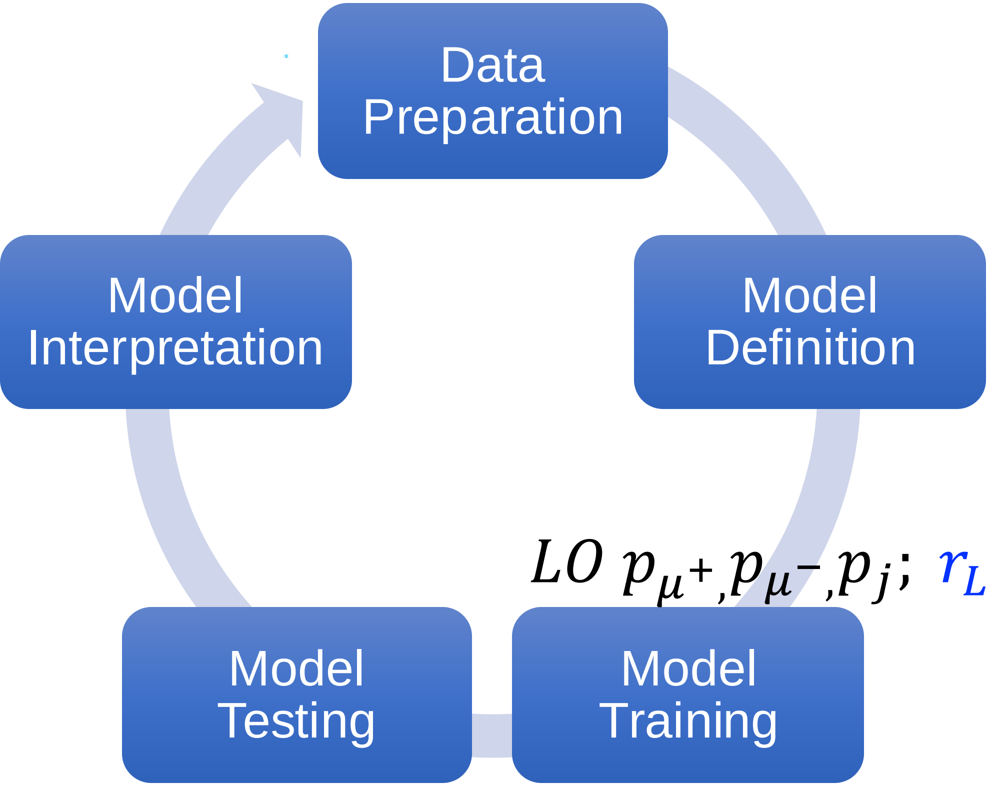
$LO\ p_{\mu^+}, p_{\mu^-}, p_j;\ \tilde{r}_L$



Working at per-cent level

NB: the NN-training stage is tailored to the specific choice of polarisation frame.

Polarization tagging LO + PS



Data Preparation

Model Definition

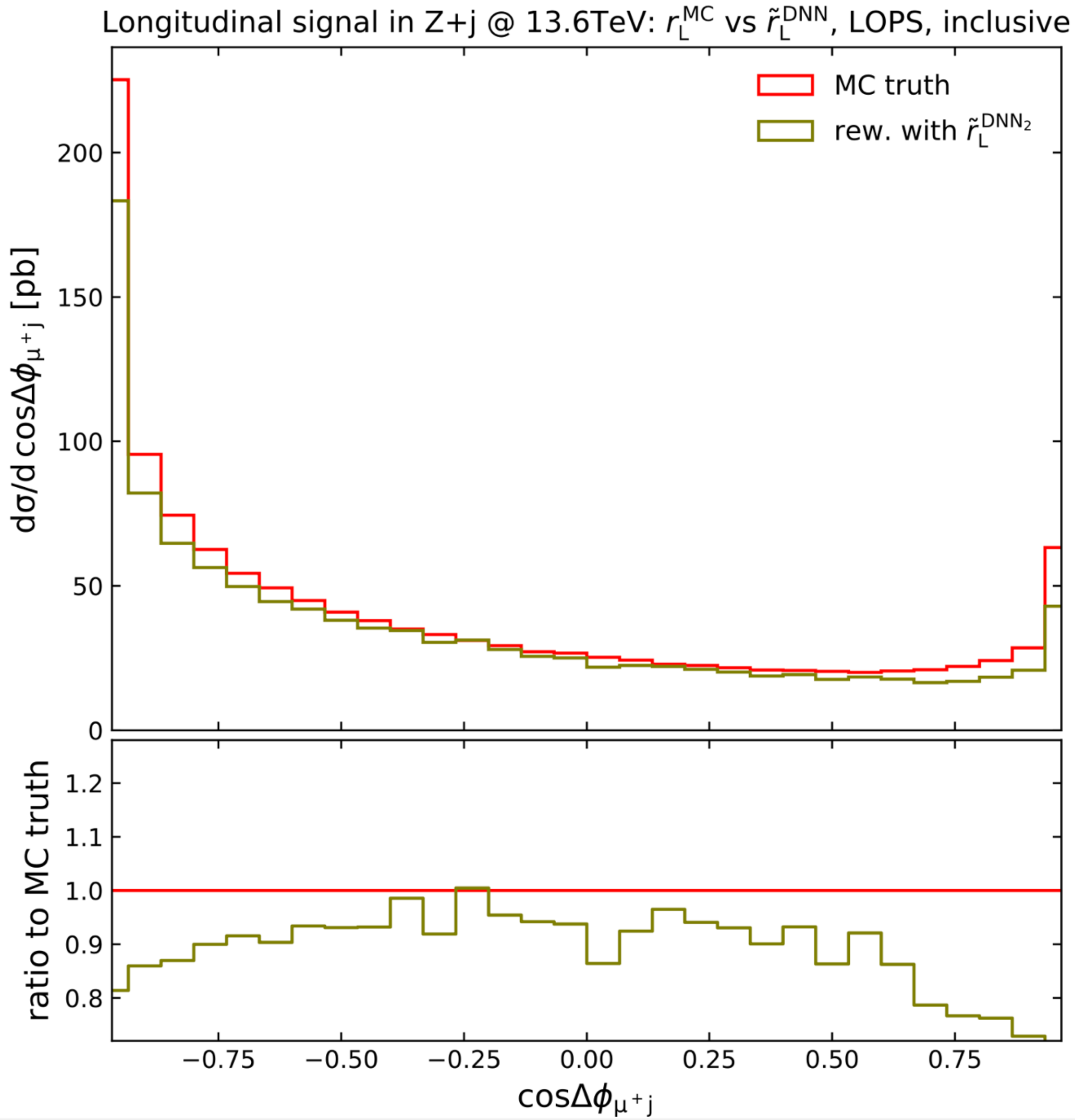
Model Training

Model Testing

Model Interpretation

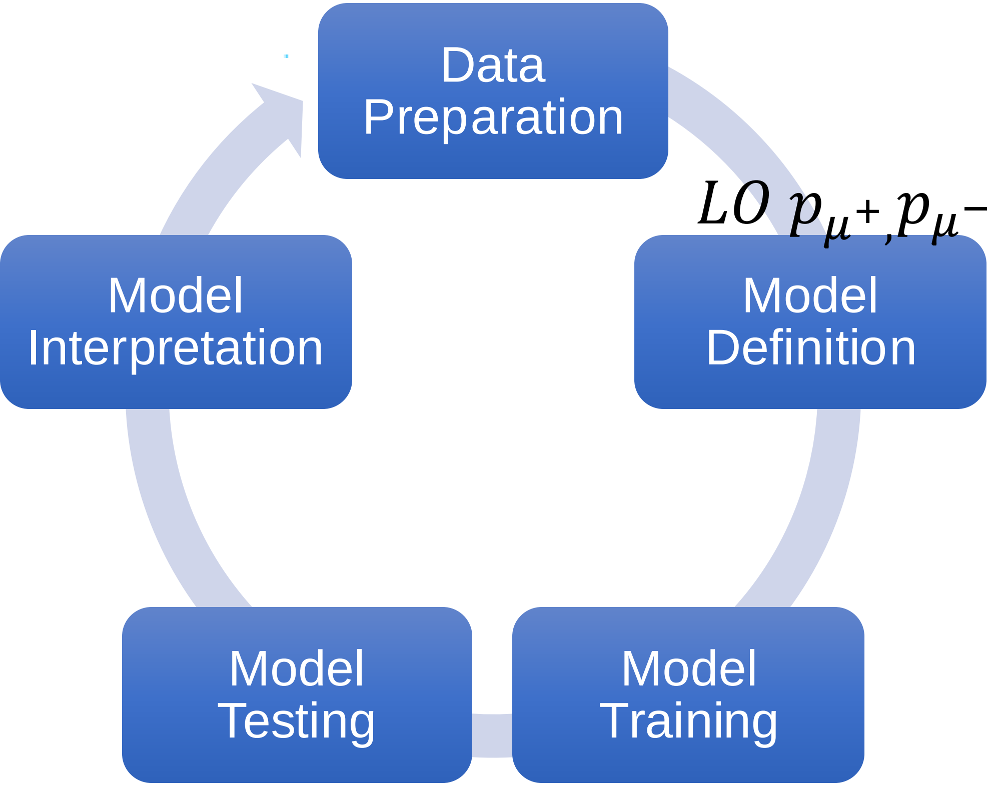
$LO\ p_{\mu^+}, p_{\mu^-}, p_j;\ r_L$

$LO + PS\ p_{\mu^+}, p_{\mu^-}, p_j;\ \widetilde{r_L}$



Failing! (describing PS corrections instead of polarisation) → Retraining with r_L computed before PS!

Polarization tagging – Warm-start



Data Preparation

Model Definition

Model Training

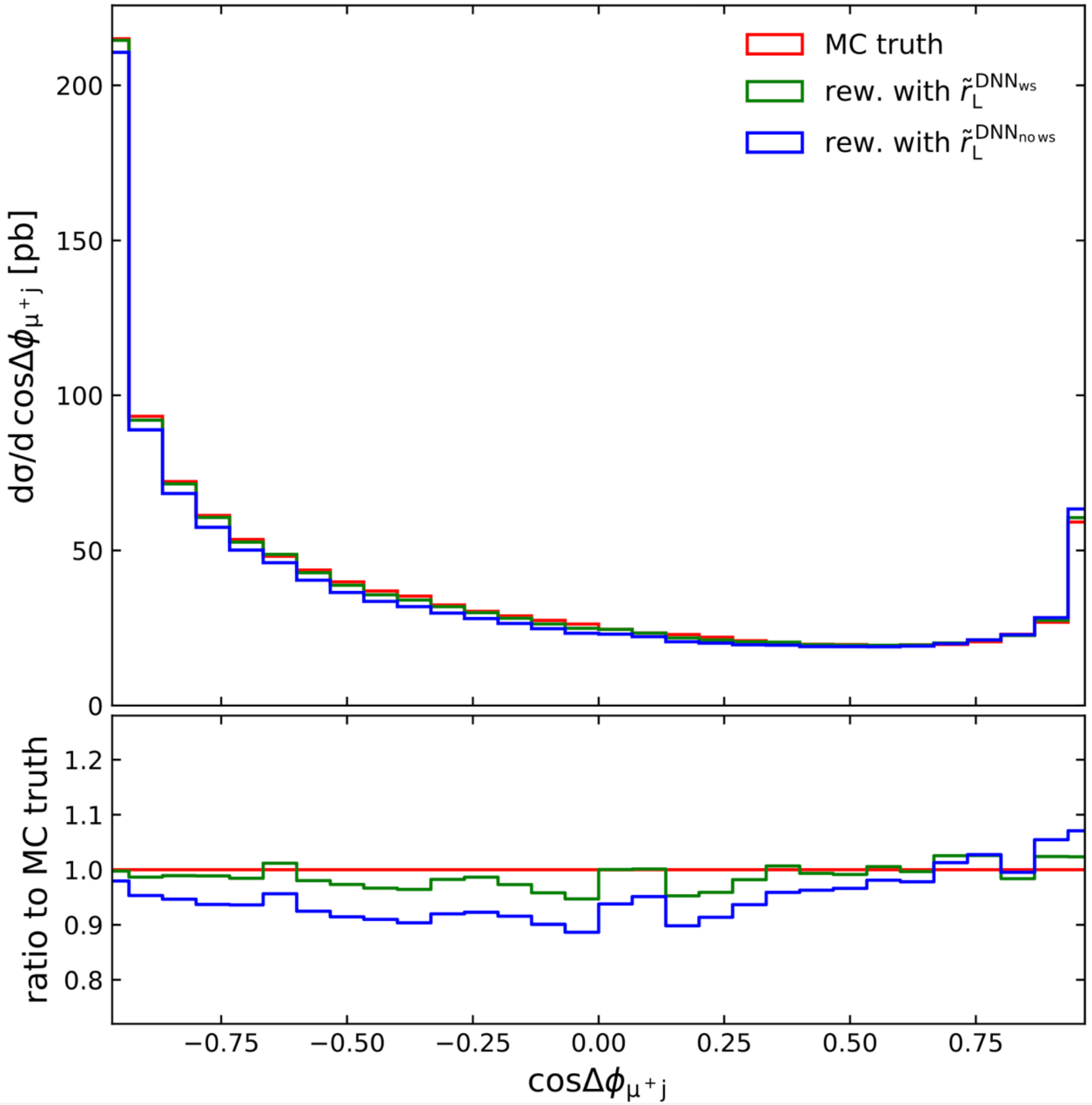
Model Testing

Model Interpretation

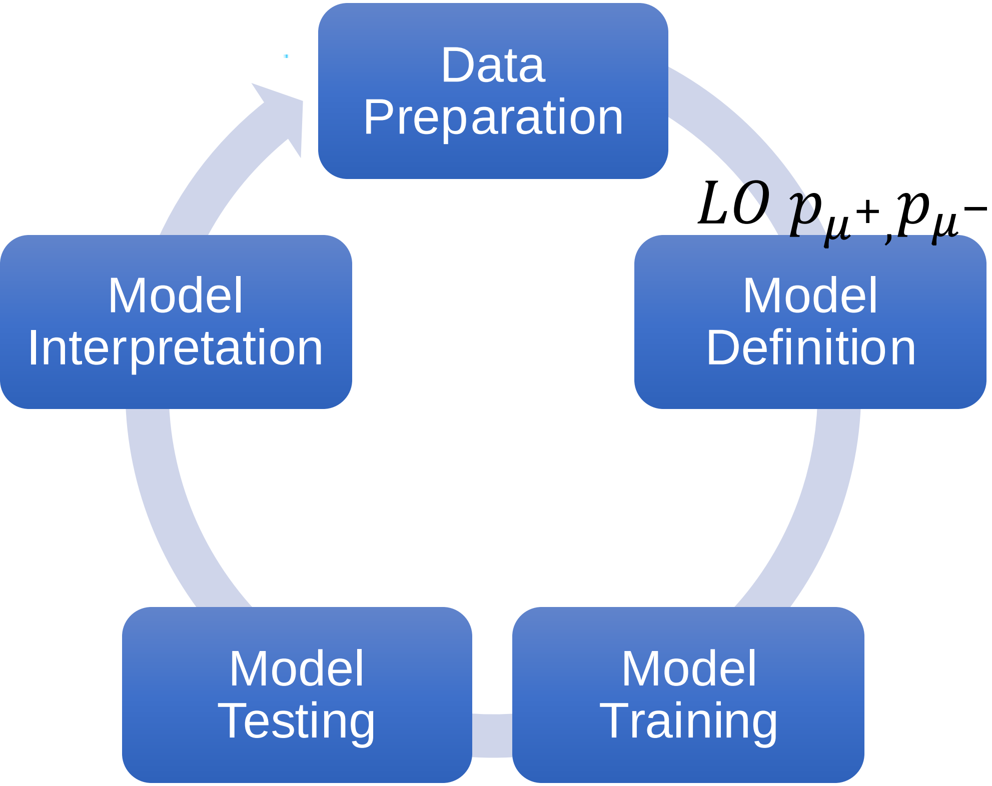
$LO\ p_{\mu^+},p_{\mu^-},p_j;\ r_L$

$LO + PS\ p_{\mu^+},p_{\mu^-},p_j;\ \tilde{r}_L$

Longitudinal signal in Z+j @ 13.6TeV: r_L^{MC} vs $\tilde{r}_L^{DNN}$, LOPS, inclusive



Polarization tagging – Warm-start



$LO\ p_{\mu^+}, p_{\mu^-}, p_j;\ r_L$

$LO + PS\ p_{\mu^+}, p_{\mu^-}, p_j;\ \widetilde{r_L}$

- Warmup training gives better results
- Initial conditions of the LO+PS learning is set by LO learning
- Satisfactory performance for integrated fractions (1% level), angular observables and bulk of p_T ones (dedicated training needed in least-populated phase-space regions).
- LO+PS is better reproduced than LO (less r_L˜< 0)

- **Problem Definition**
- **ML & Polarization Tagging**
- **Results**
- **Summary & Outlook**

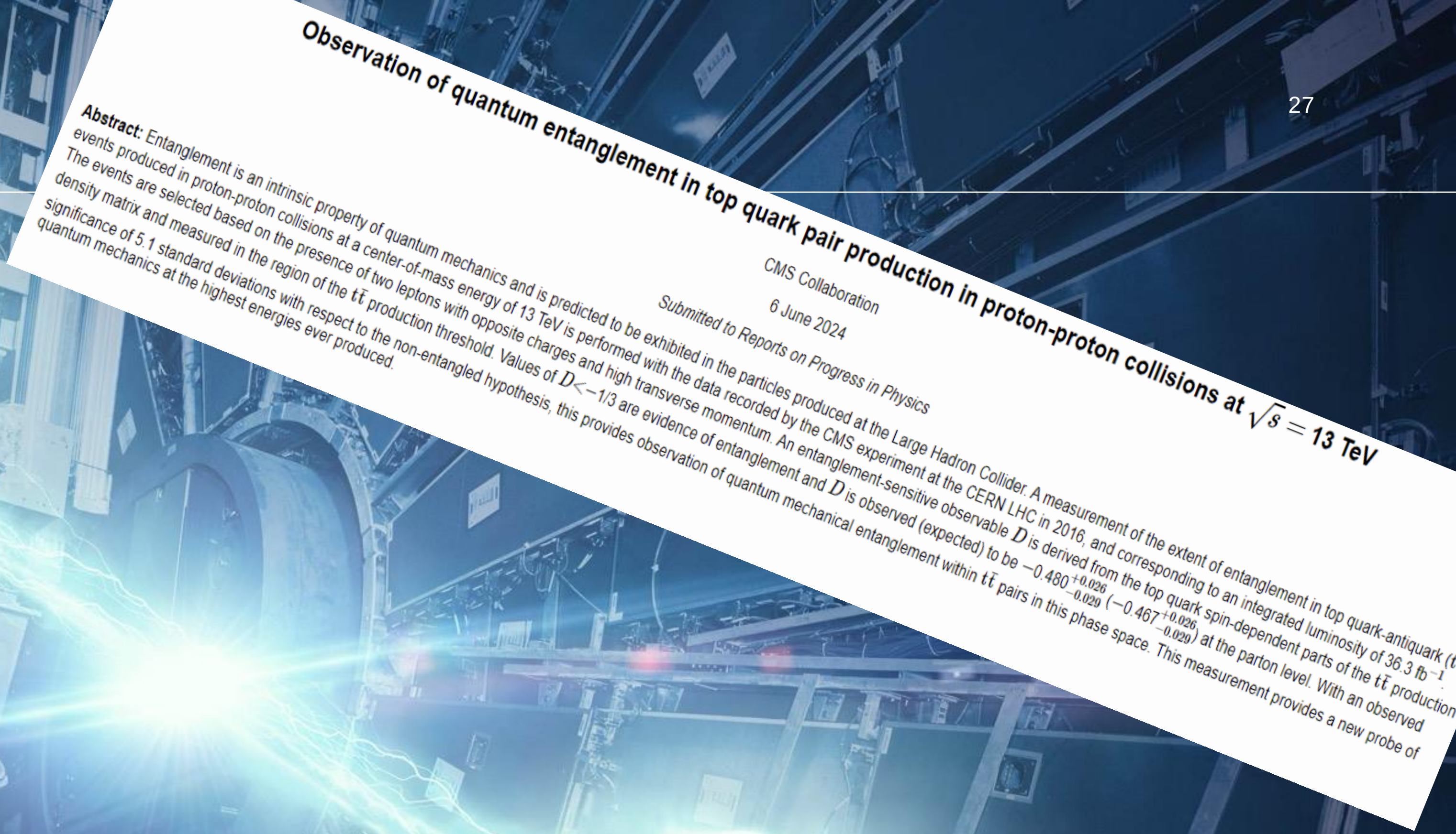
Discussion and Outlook

- Propose an original method to extract polarisation fractions using the amplitude
- The method can be applied to any physics problem that can be cast in one single, bounded, ratio (in the present application rL) which can be approximately reconstructed using incomplete information thanks to NN methods
- Applicable at different energy?
- Higher orders

  → Possible reweighting with virtual and real amplitudes
- Use physics-informed approaches, enforcing symmetries conservation?

  BUT there is no guarantees

$$\begin{tikzcd}
SO(3;1)^+ \arrow[r, "\Pi"] & GL(V) \\
\mathfrak{so}(3;1) \arrow[u, "\exp"] \arrow[r, "\pi"'] & \mathfrak{gl}(V) \arrow[u, "\exp"']
\end{tikzcd}$$



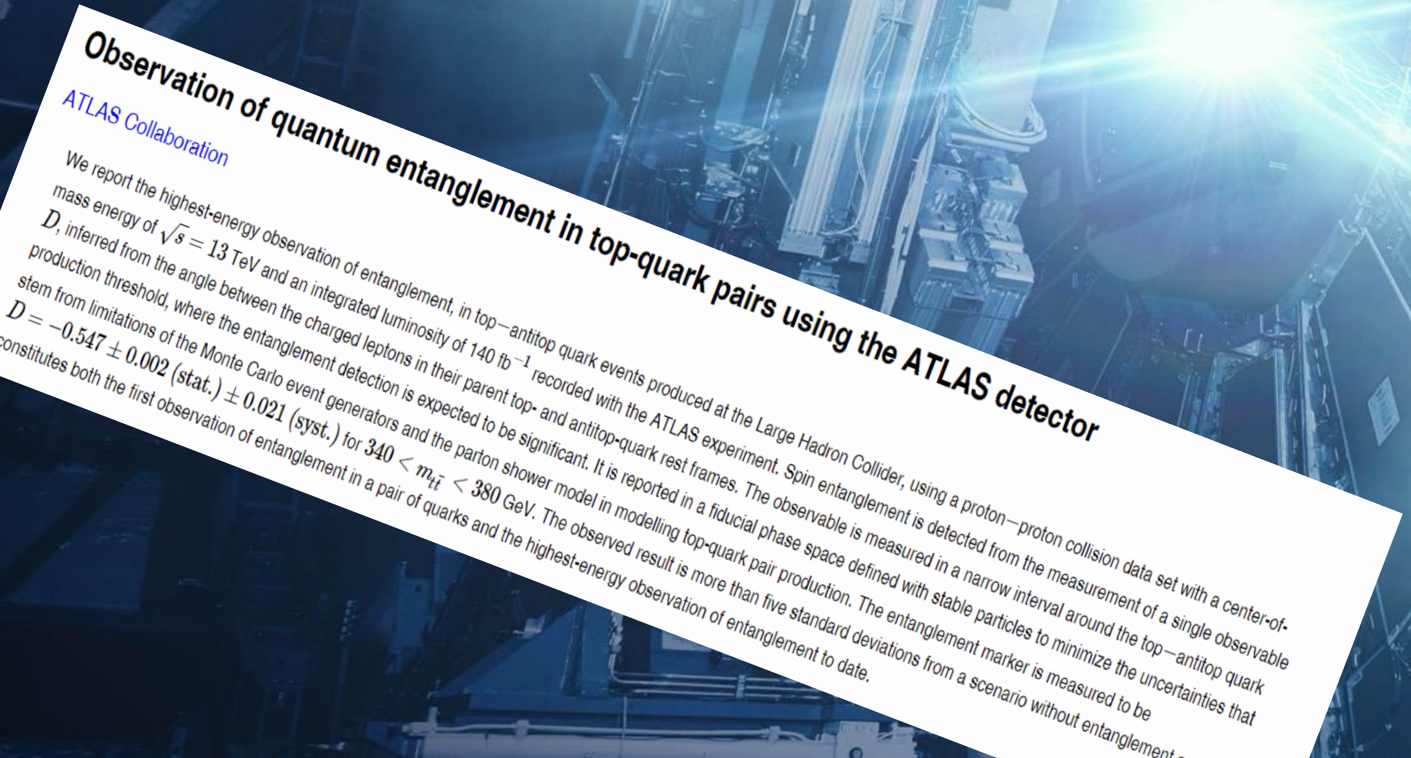
Observation of quantum entanglement in top quark pair production in proton-proton collisions at $\sqrt{s} = 13$ TeV

CMS Collaboration

6 June 2024

Submitted to Reports on Progress in Physics

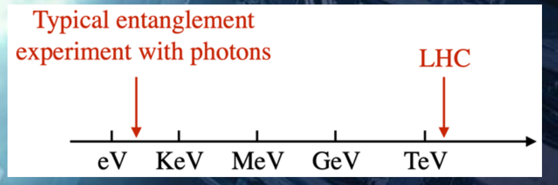
Abstract: Entanglement is an intrinsic property of quantum mechanics and is predicted to be exhibited in the particles produced at the Large Hadron Collider. A measurement of the extent of entanglement in top quark-antiquark (t events produced in proton-proton collisions at a center-of-mass energy of 13 TeV is performed with the data recorded by the CMS experiment at the CERN LHC in 2016, and corresponding to an integrated luminosity of 36.3 fb^{-1}. The events are selected based on the presence of two leptons with opposite charges and high transverse momentum. An entanglement-sensitive observable D is derived from the top quark spin-dependent parts of the $t\bar{t}$ production density matrix and measured in the region of the $t\bar{t}$ production threshold. Values of $D<-1/3$ are evidence of entanglement and D is observed (expected) to be $-0.480^{+0.026}_{-0.029}$ $(-0.467^{+0.026}_{-0.029})$ at the parton level. With an observed significance of 5.1 standard deviations with respect to the non-entangled hypothesis, this provides observation of quantum mechanical entanglement within $t\bar{t}$ pairs in this phase space. This measurement provides a new probe of quantum mechanics at the highest energies ever produced.



Observation of quantum entanglement in top-quark pairs using the ATLAS detector

ATLAS Collaboration

We report the highest-energy observation of entanglement, in top−antitop quark events produced at the Large Hadron Collider, using a proton−proton collision data set with a center-of-mass energy of $\sqrt{s} = 13$ TeV and an integrated luminosity of 140 fb^{-1} recorded with the ATLAS experiment. Spin entanglement is detected from the measurement of a single observable D, inferred from the angle between the charged leptons in their parent top- and antitop-quark rest frames. The observable is measured in a narrow interval around the top−antitop quark production threshold, where the entanglement detection is expected to be significant. It is reported in a fiducial phase space defined with stable particles to minimize the uncertainties that stem from limitations of the Monte Carlo event generators and the parton shower model in modelling top-quark pair production. The entanglement marker is measured to be $D = -0.547 \pm 0.002\ (\text{stat.}) \pm 0.021\ (\text{syst.})$ for $340 < m_{t\bar{t}} < 380$ GeV. The observed result is more than five standard deviations from a scenario without entanglement constitutes both the first observation of entanglement in a pair of quarks and the highest-energy observation of entanglement to date.



CERN QTI

https://quantum.cern/