A $W^{\pm}$ polarization analyzer from Deep Neural Networks

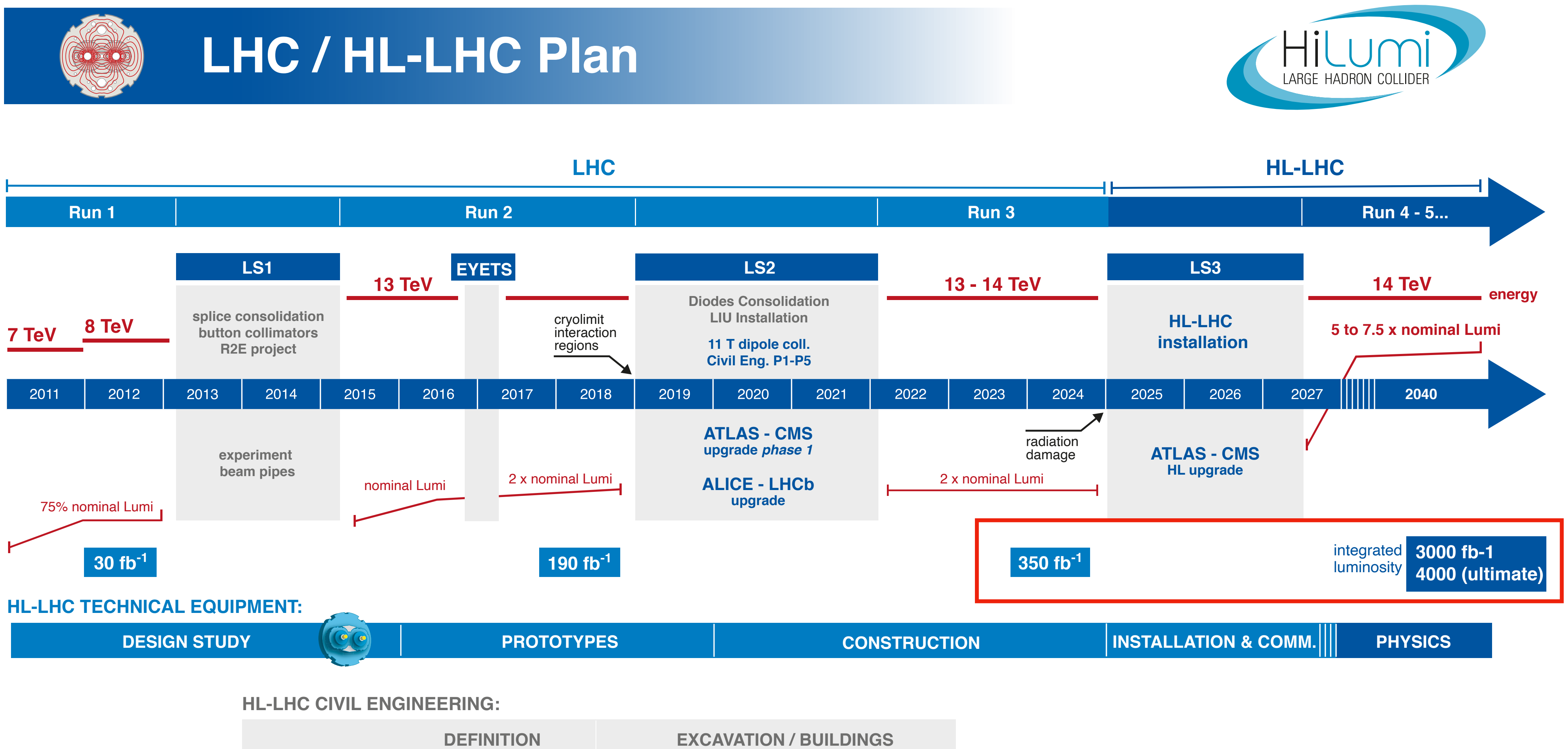
Taegyun Kim
Research Advisor: Dr. Adam Martin
Department of Physics, University of Notre Dame

arXiv:2102.05124

5/24/2021 Pheno

Introduction

Where are we now?



https://project-hl-lhc-industry.web.cern.ch/content/project-schedule

- Entering HL-LHC : bring out small number of event signals
- Precision testing to find potential BSM signatures
- This research is about building a tool and show possibility

Theoretical Motivation

Massive vector boson final states

$$p\,p \to W^{\pm}\,W^{\mp}$$

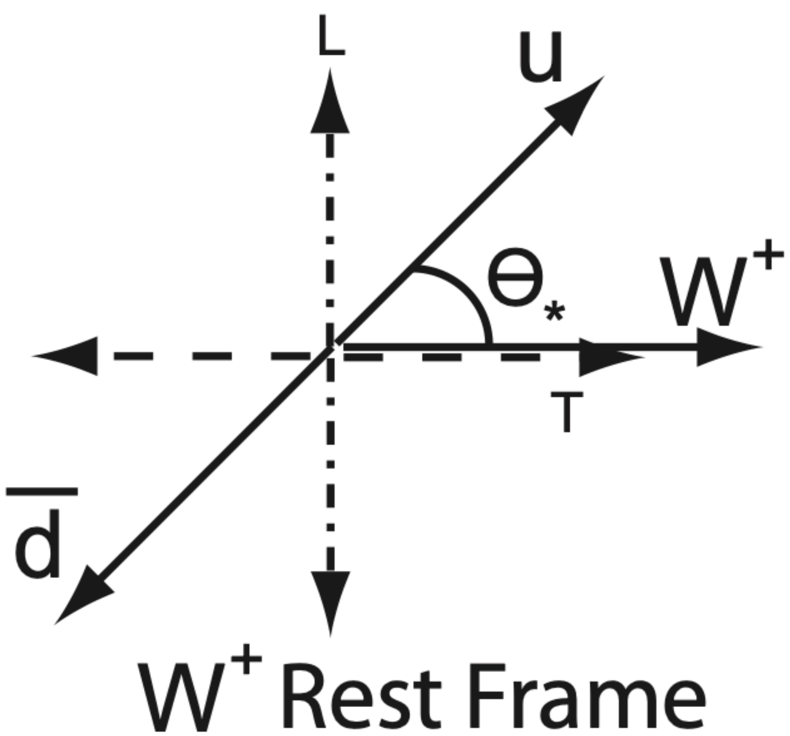
$$p\,p \to W^{\pm}\,Z$$

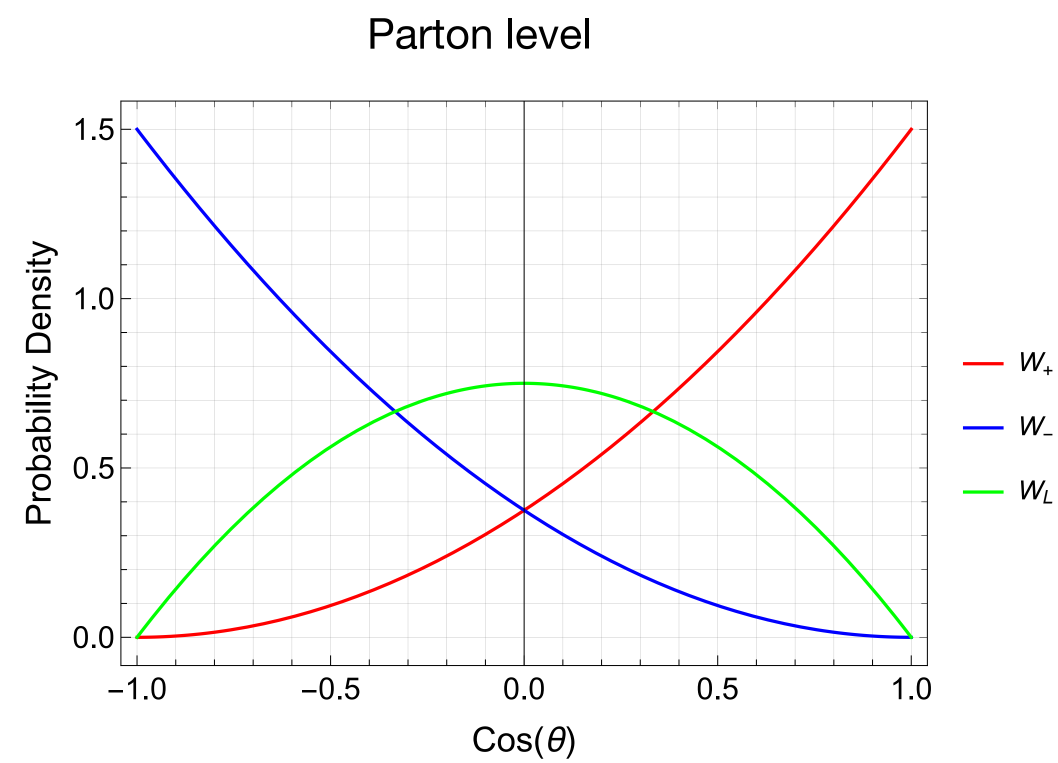
$$p\,p \to Z\,Z$$

- Indirect approach of checking SM : polarization searches
  - Longitudinal vs. Transverse
- SM can predict polarization fraction
- Longitudinal polarization is sensitive to EWSB
- Some SMEFT operators can affect longitudinal fraction of a process

W polarization

Decay of W

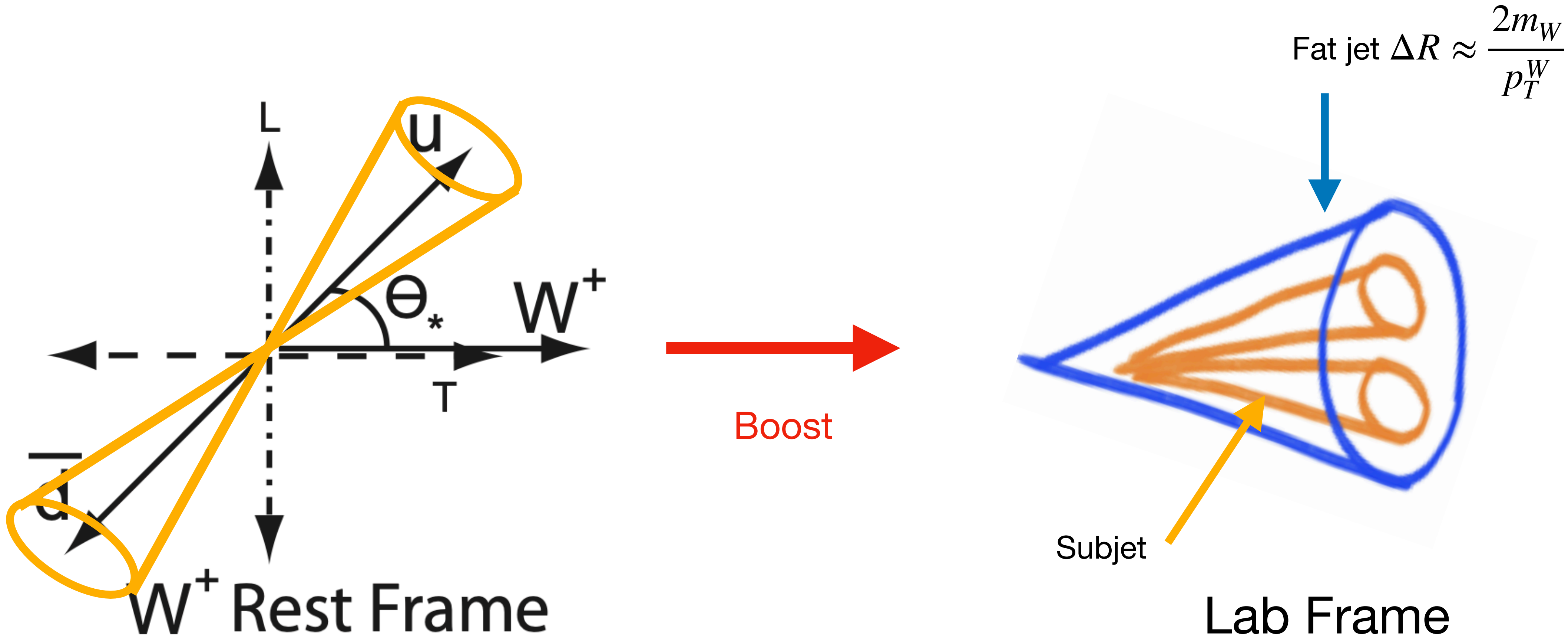




- There is a limitation in leptonically decaying W
- Since W only interacts to the left handed particles, each polarization has distinct angular distribution
- Due to the deviation, it is possible to measure polarization fraction for diboson final states
- Large overlap in parton level distribution may suppress even by event tagging

Boosted W Jet

Decay of W



- Quark becomes QCD jet
- Due to the boost, collimation of the jet deduces the angular distribution signature
- Possible subjet signature
- After boost $\theta^* \rightarrow$ opening angle (sensitive to pT)
- At extreme high p_T^W, subjet signature can disappear

Machine Learning Motivation

Machine Learning in HEP

- The current most frequently used machine learning algorithm: Boosted Decision Tree (BDT) and Neural Network (NN)
- Major usage
  - Classification : PID, event identification
  - Regression : predict particle energy

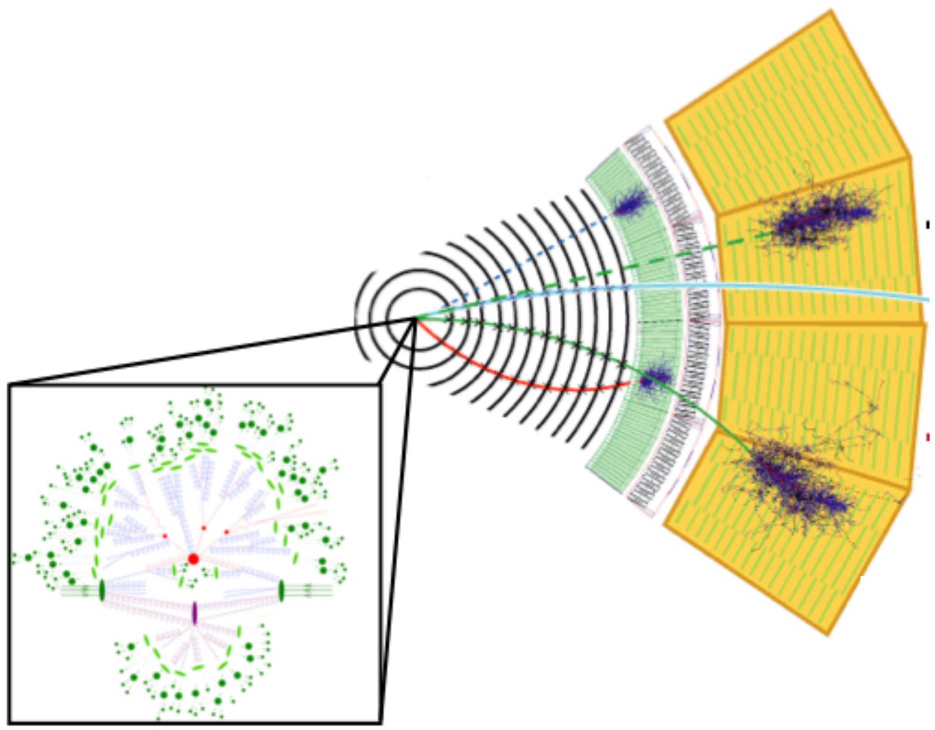
Our interest

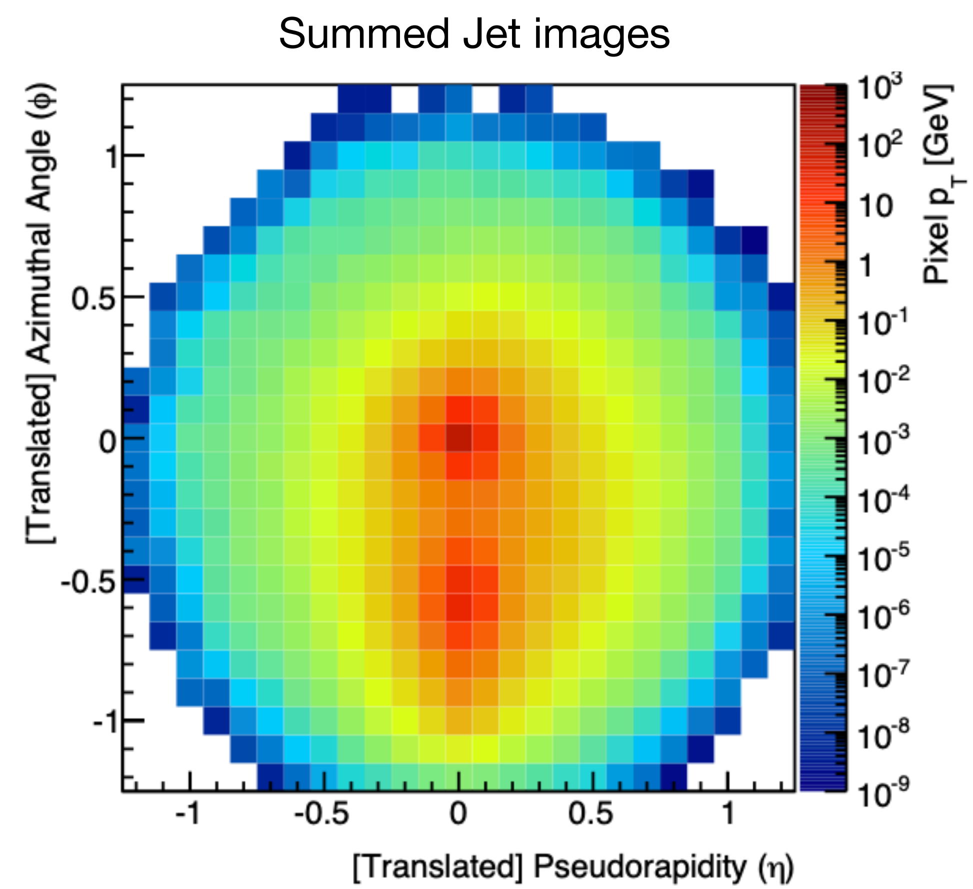
- Recent researches on: quark vs. gluon, QCD vs. top, W vs. QCD

ML(NN) technique to distinguish hadronic W's polarization to test on $W^{\pm}Z$ final state

Jet as an Image from Collider

Adjust for HEP using jet image



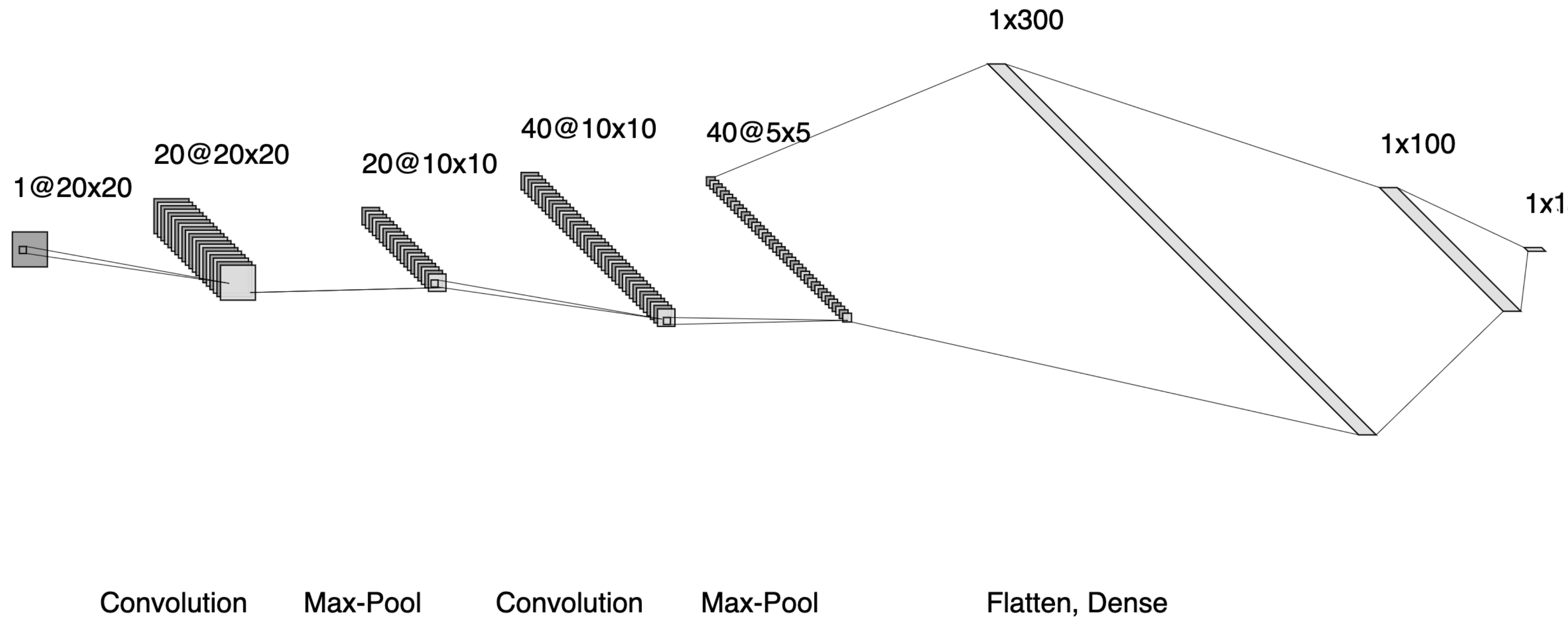


L. de Oliveira, M. Kagan, L. Mackey, B. Nachman, and A. Schwartzman, “Jet-images – deep learning edition,”

- In collider, images are created from outgoing particles
- Particles are plotted on pixelized $\eta - \phi$ plane and their color is determined from p_T

Convolutional Neural Network (CNN)

Image recognition



- Ordinary CNN structure : Convolution - Flatten - Dense

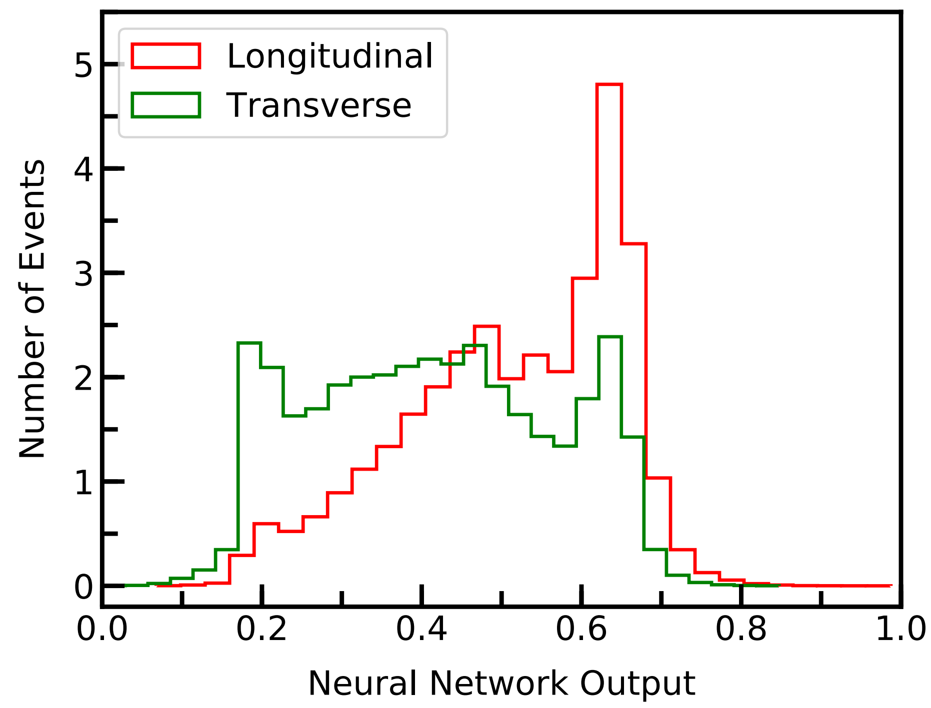
- The network is trained with simulated events **(MadGraph** + **Pythia** + **Delphes)** of boosted longitudinal and transverse W's respectively for tagging purposes

- Depending on p_T^W, images are separated into 2 bins: [200,300] and [400,500] since for fat jet, $\Delta R \approx \dfrac{2m_W}{p_T^W}$

Testing on SM

Longitudinal fraction (f_L)

[200, 300]



- Checking distribution can tell us how good the separation between two polarization
- Inhibits potential event by event tagging because of large overlap

- In order to apply for testing, we measure f_L of randomly selected events
- Test on WZ final state

SM

p_T range	$\sigma(pp \to W^{\pm}(jj)Z(\ell\ell))$ (fb)	truth $\sigma_L/\sigma_{\text{tot}}$	predicted f_L
$200\,\text{GeV} \leq p_T \leq 300\,\text{GeV}$	6.67	0.265	0.259 ± 0.013
$400\,\text{GeV} \leq p_T \leq 500\,\text{GeV}$	0.35	0.304	0.300 ± 0.033

SMEFT in Diboson Final States

SMEFT intro

- $\mathscr{L}_{SMEFT} = \mathscr{L}_{SM} + \sum_{D>4}^{\inf} \frac{1}{\Lambda^{D-4}} c_j^{(D)} \mathcal{O}_j^{(D)}$

- SMEFT extends the SM Lagrangian by gauge invariant higher dim (D>4) operators

- We will investigate boosted W cases

Relevant operators (SILH) for diboson final states

Da Liu, Lian-Tao Wang [arXiv: 1804.08688v1]

$$\boxed{\mathcal{O}_W = \frac{ig}{2}\left(H^\dagger \sigma^a \overleftrightarrow{D}^\mu H\right) D^\nu W^a_{\mu\nu}}$$

Longitudinal

$$\mathcal{O}_B = \left(H^\dagger \sigma^a \overleftrightarrow{D}^\mu H\right) \partial^\nu B_{\mu\nu}$$

$$\mathcal{O}_{2W} = -\frac{1}{2} D^\mu W^a_{\mu\nu} D_\rho W^{a\rho\nu}$$

$$\boxed{\mathcal{O}_{3W} = \frac{1}{3!} g \epsilon_{abc} W_\mu^{a\nu} W^b_{\nu\rho} W^{c\rho\mu}}$$

Transverse

$$\mathcal{O}_{HW} = ig\left(D^\mu H\right)^\dagger \sigma^a \left(D^\nu H\right) W^a_{\mu\nu}$$

$$\mathcal{O}_{HW} = ig'\left(D^\mu H\right)^\dagger \left(D^\nu H\right) B_{\mu\nu}$$

Possible Scenarios with SMEFT

SM

p_T range	$\sigma(pp \to W^{\pm}(jj)Z(\ell\ell))$ (fb)	truth $\sigma_L/\sigma_{\text{tot}}$	predicted f_L
200 GeV $\leq p_T \leq$ 300 GeV	6.67	0.265	0.259 ± 0.013
400 GeV $\leq p_T \leq$ 500 GeV	0.35	0.304	0.300 ± 0.033

1. Shift longitudinal fraction with cross section shift

	p_T range	$\sigma(pp \to W^{\pm}Z)$ (fb)	truth σ_L/σ_{tot}	predicted f_L
O_W	200 GeV $\leq p_T \leq$ 300 GeV	6.93	0.311	0.297 ± 0.010
	400 GeV $\leq p_T \leq$ 500 GeV	0.42	0.439	0.391 ± 0.033
O_{3W}	200 GeV $\leq p_T \leq$ 300 GeV	6.58	0.258	0.254 ± 0.011
	400 GeV $\leq p_T \leq$ 500 GeV	0.50	0.198	0.181 ± 0.043

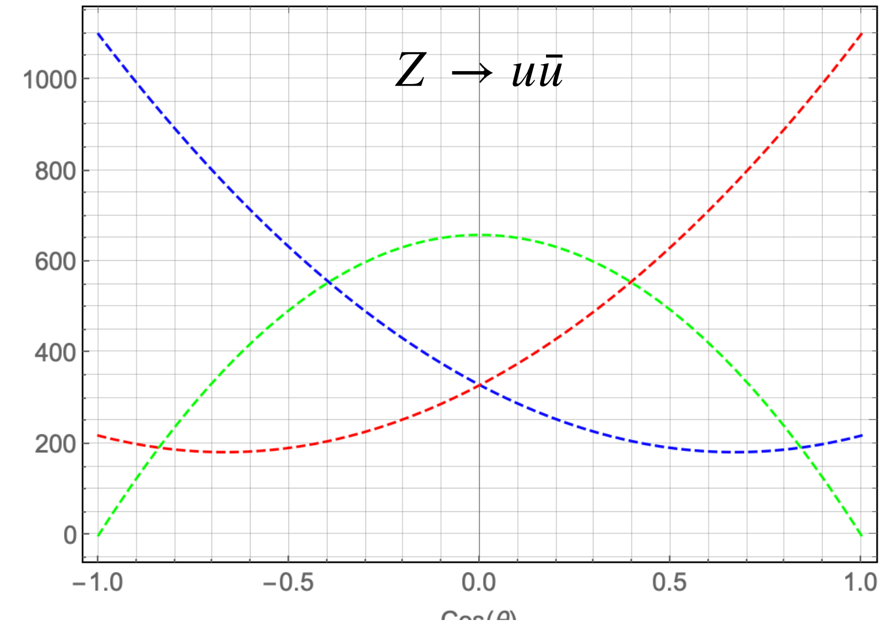
2. Shift longitudinal fraction without cross section shift

SM + $\mathcal{O}_W$ + $\mathcal{O}_{3W}$

p_T range	$\sigma(pp \to W^{\pm}Z)$ (fb)	truth σ_L/σ_{tot}	predicted f_L
200 GeV $\leq p_T \leq$ 300 GeV	6.68	0.202	0.207 ± 0.011
400 GeV $\leq p_T \leq$ 400 GeV	0.34	0.285	0.282 ± 0.044

Conclusion/Discussion

- Simple CNN can be used to tag $W^{\pm}$ polarization though event by event tagging is suppressed
- Ensemble analysis using network's output average values can help to predict f_L
- Network prediction can catch small f_L deviations originated from dim 6 operators
- If cross section changes, polarization measurement can clear out degeneracies between EFT operators
- Possible applicability on Z jets
- Potential limits
  - $W^{\pm}$ vs. Z vs. QCD is not perfectly separable
  - Cuts that can cause polarization interference



Thank you

References

(1) Aaboud, M., Aad, G., Abbott, B., Abdinov, O. et al. (2019). Measurement of $W^{\pm}Z$ production cross sections and gauge boson polarization in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector. *The European Physical Journal C*, 79(6).

(2) D. Liu and L.-T. Wang, “Prospects for precision measurement of diboson processes in the semileptonic decay channel in future LHC runs,”Physical Review D 99 (Mar, 2019) . http://dx.doi.org/10.1103/PhysRevD.99.055001.

(3) Carleo, Giuseppe et al. "Machine learning and the physical sciences". Reviews of Modern Physics 91. 4(2019).

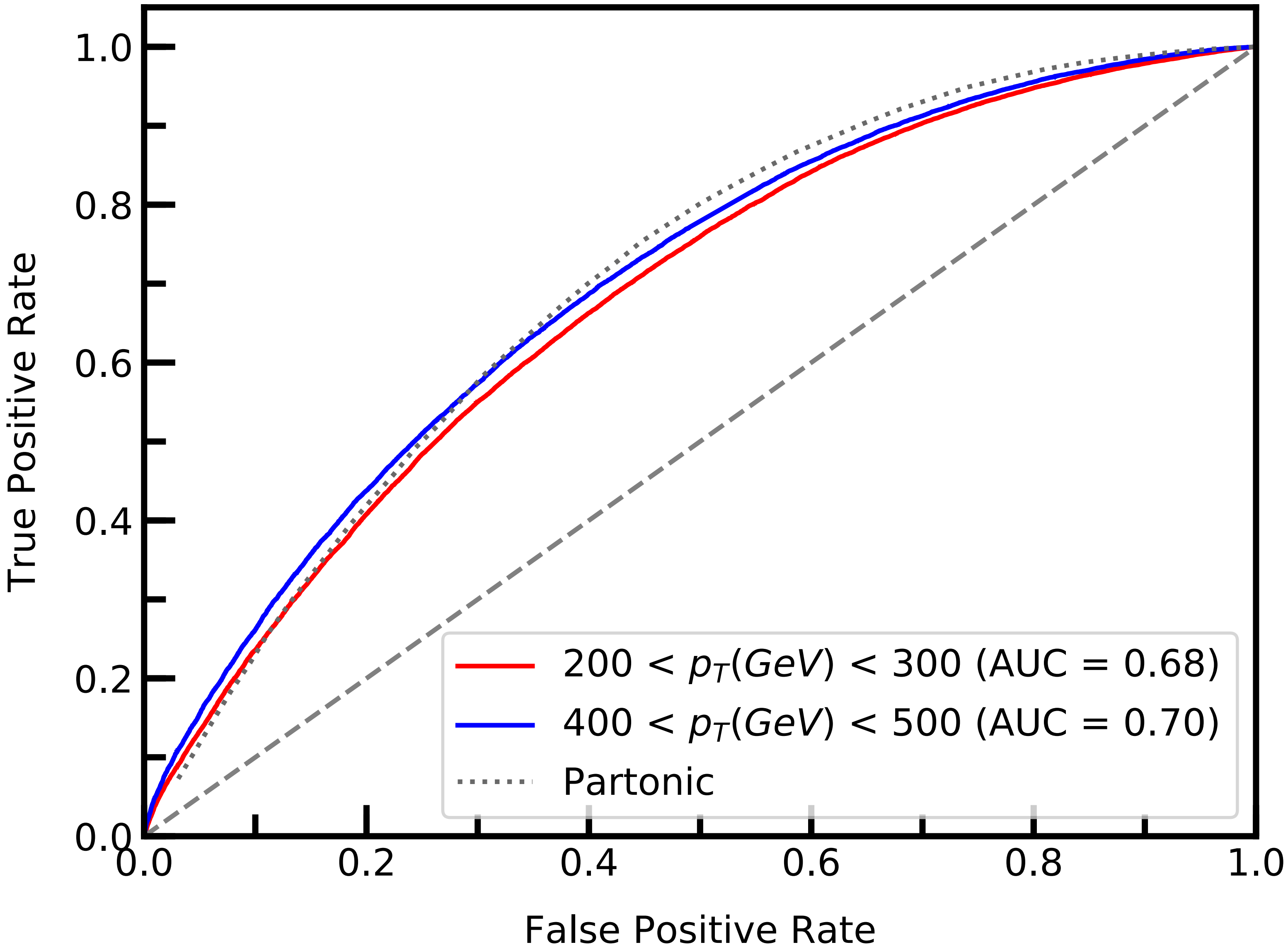
(4) Stirling, W. J. et al. "Electroweak gauge boson polarisation at the LHC". Journal of High Energy Physics 2012. 7(2012).

(5) “scikit-hep/pyjet: 1.6.0 (version 1.6.0),”

Backup slides

Training Quality

Distribution check

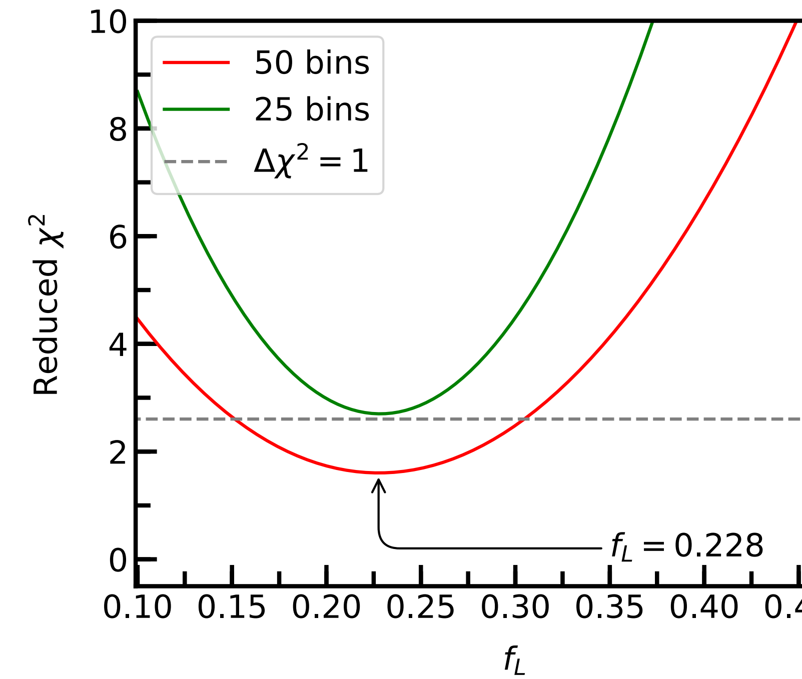


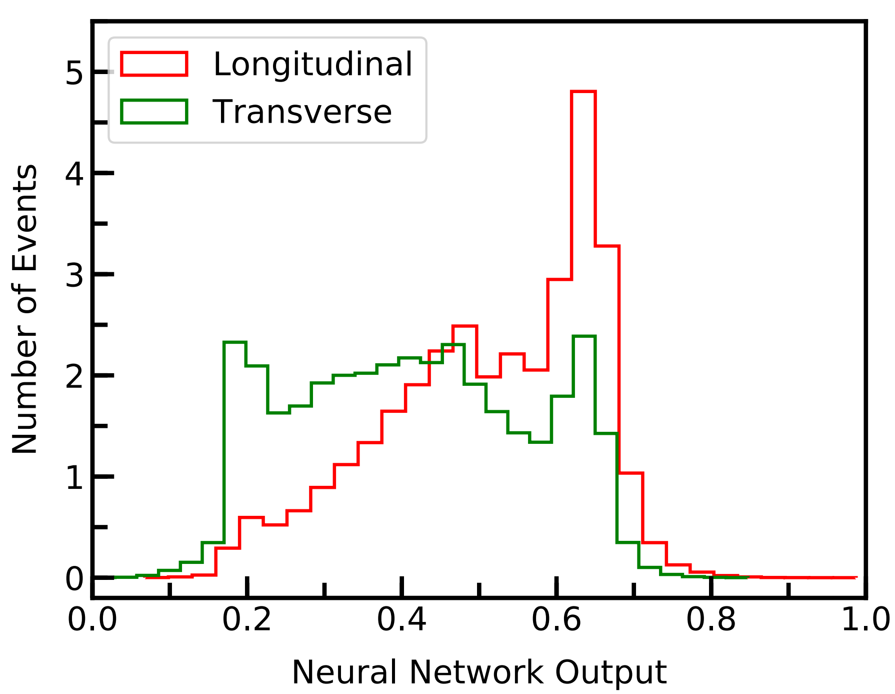
- Checking distribution can tell us how good the separation between Logi and trans is.
- Inhibits potential event by event tagging since accuracy is ~ 60%
- Ensemble distribution checking to find longitudinal fraction (f_L)

Simpler Method

Network output average method

- Template fitting method depends on finding "sweet spot" for f_L
  - number of bins
  - find minimum $\chi^2(f_L)$
- Simplify by treating output distribution as probability distribution





$$\int x dx \left(D_u(x) = f_L D_L(x) + (1 - f_L) D_T(x) \right)$$

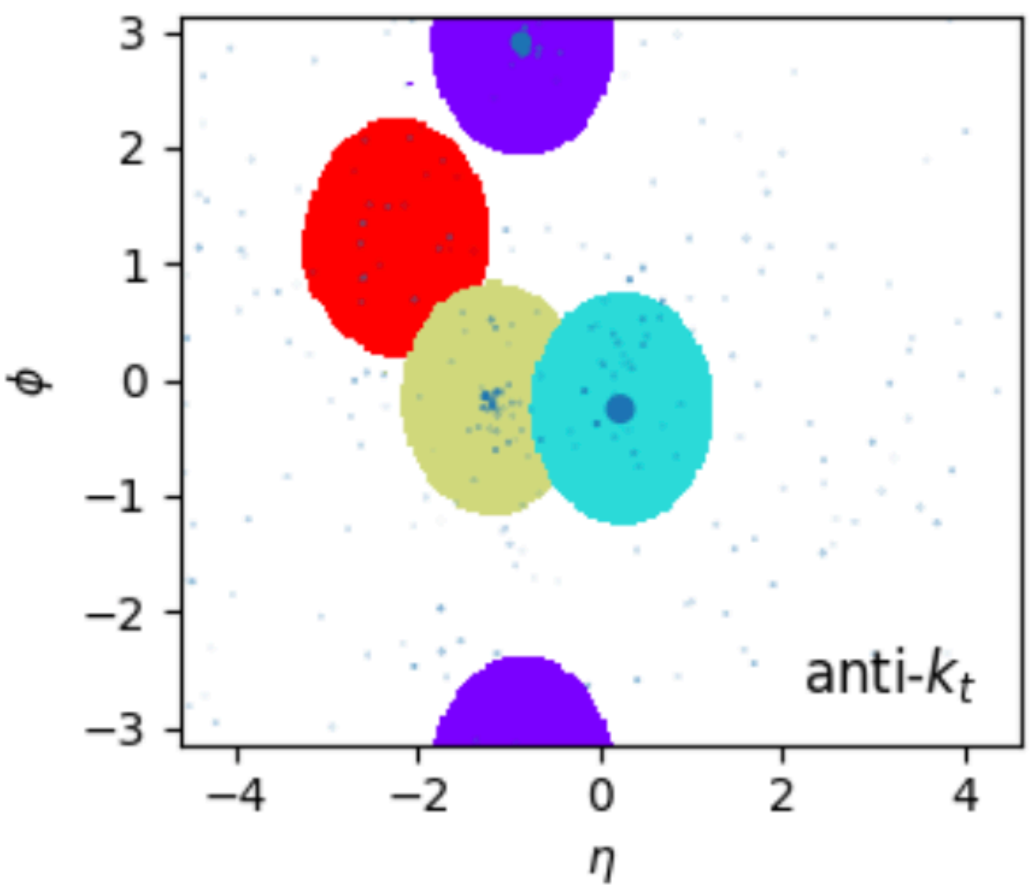
$$\langle x_u \rangle = f_L \langle x_L \rangle + (1 - f_L) \langle x_T \rangle$$

$$f_L = \frac{\langle x_u \rangle - \langle x_T \rangle}{\langle x_L \rangle - \langle x_T \rangle}$$

Confirmed that both yield the same result

Jet Images

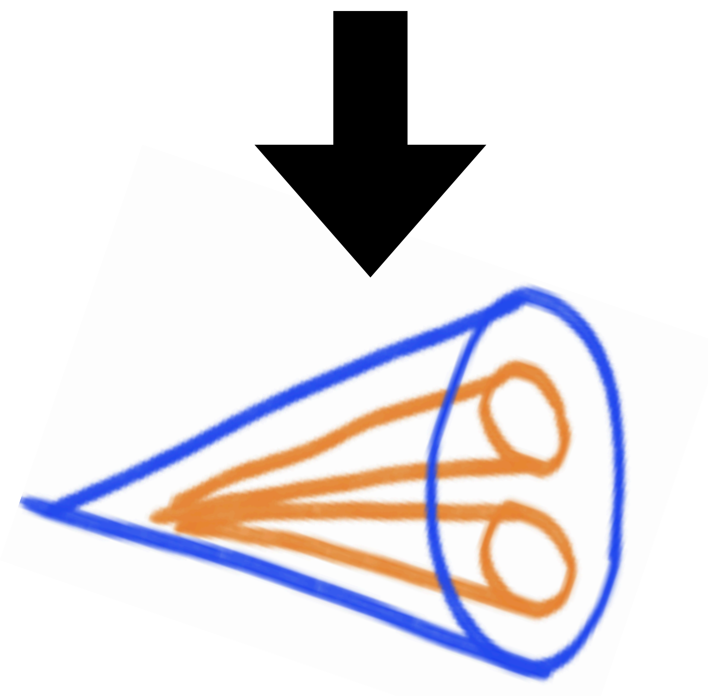
Network friendly form



https://github.com/scikit-hep/pyjet

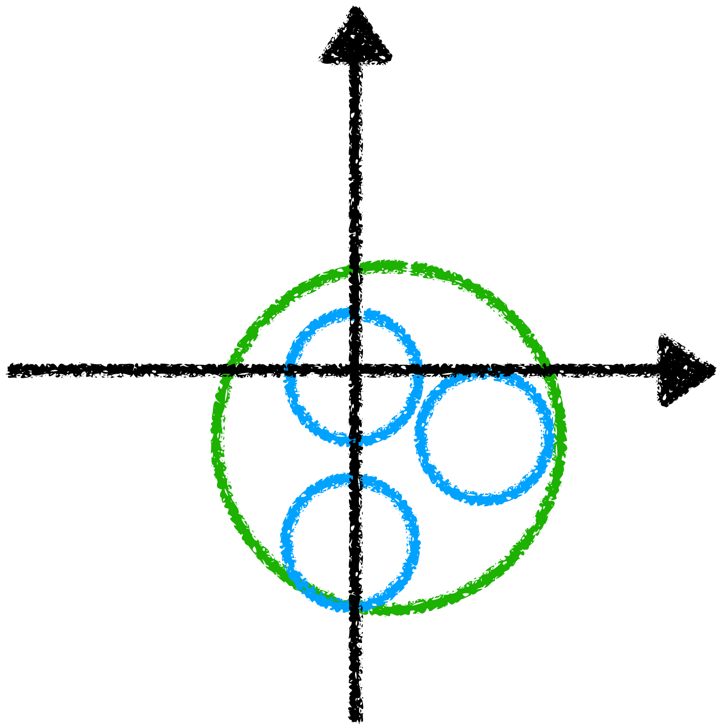
Bring out subjet signature

1. Identify jet with clustering algorithm
2. Check if clustered jet lies under p_T bin range
3. Select jets with correct angular position
4. Recluster to identify subjets



Jet Images

Network friendly form

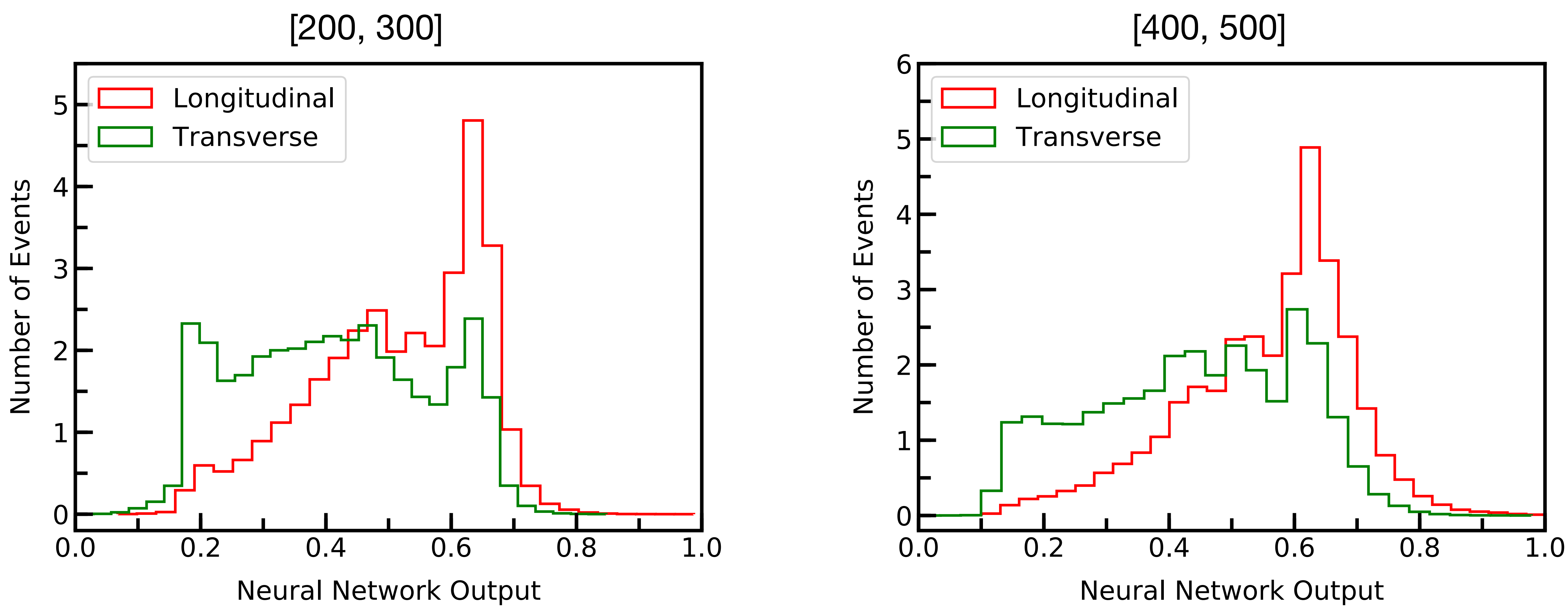


Reduce image discrepancies by putting into consistent orientation

1. Translate to centralize the highest p_T subjet
2. Rotate so that the second highest p_T subjet below the highest
3. Reflect
4. Pixelize
5. Normalize

Training Quality

Distribution check



- Checking distribution can tell us how good the separation between two polarization
- Inhibits potential event by event tagging because of large overlap
  - Putting decision threshold would contain large contamination
- Ensemble distribution checking to find longitudinal fraction (f_L)

Kinematic Cut Effect

W rest frame

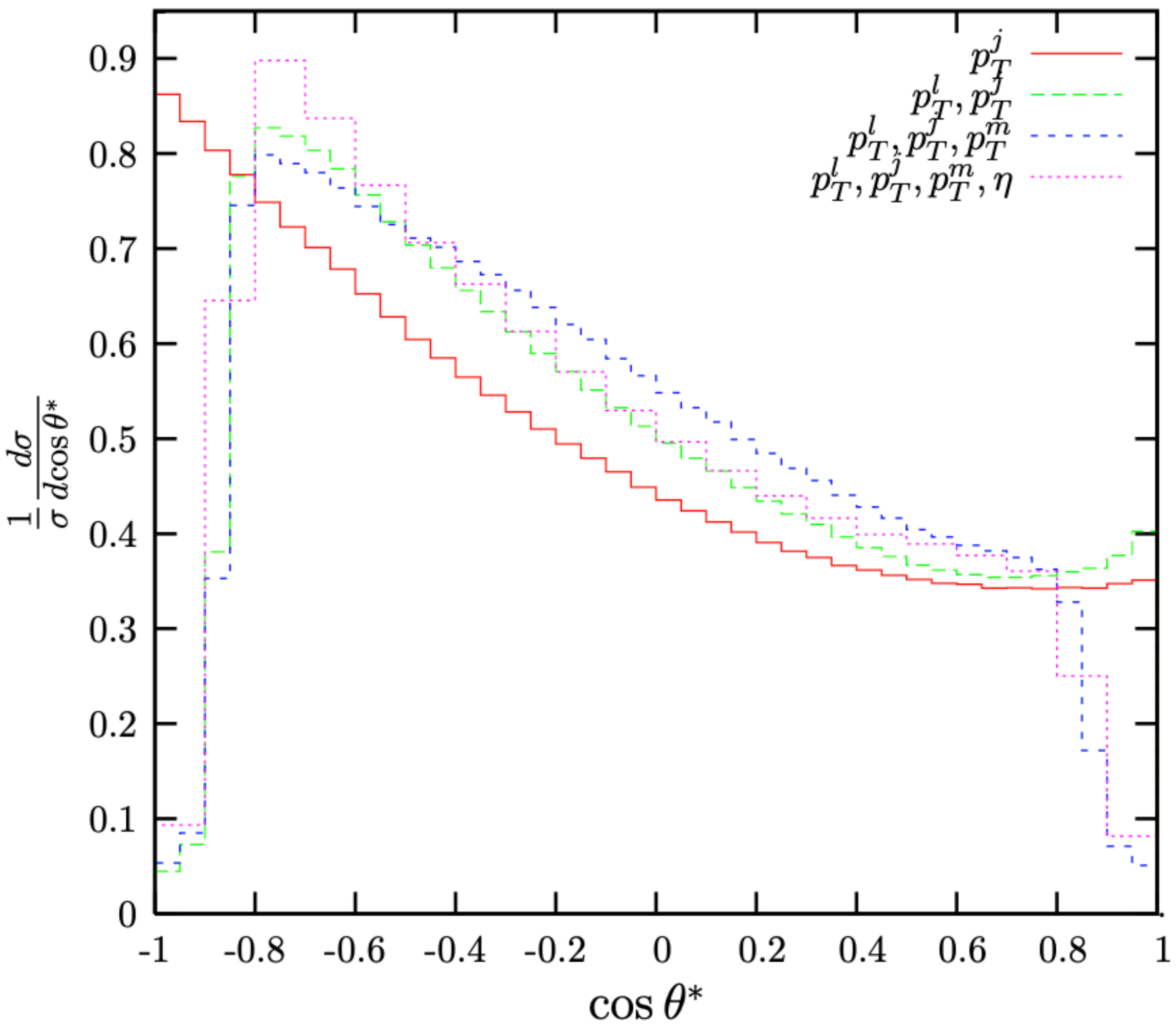
$$\frac{1}{\sigma}\frac{d\sigma}{d\cos\theta^*} = \frac{3}{8}(1-\cos\theta^*)^2 f_L + \frac{3}{8}(1+\cos\theta^*)^2 f_R + \frac{3}{4}\sin^2\theta^* f_0,$$

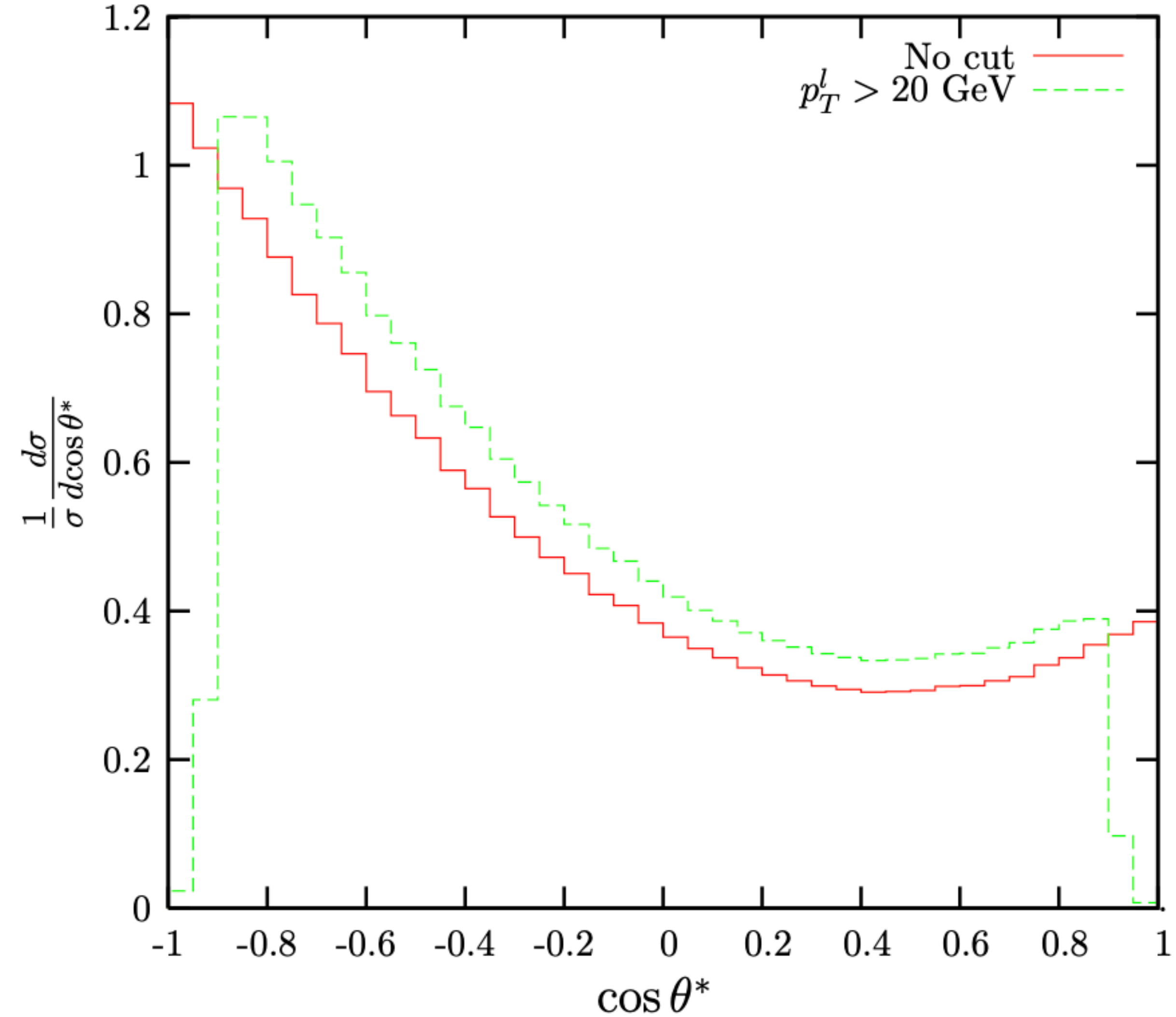
W at LHC

$$\frac{1}{\sigma}\frac{d\sigma}{d\cos\theta^* d\phi^*} = \frac{3}{16\pi}[(1+\cos^2\theta^*) + A_0\frac{1}{2}(1-3\cos^2\theta^*) + A_1\sin2\theta^*\cos\phi^* + A_2\frac{1}{2}\sin^2\theta^*\cos2\phi^* + A_3\sin\theta^*\cos\phi^* + A_4\cos\theta^*],$$

- Integrating over ϕ^* will give the same result but kinematic cut can change

Kinematic Cut Effect





Kinematic Cut Effect

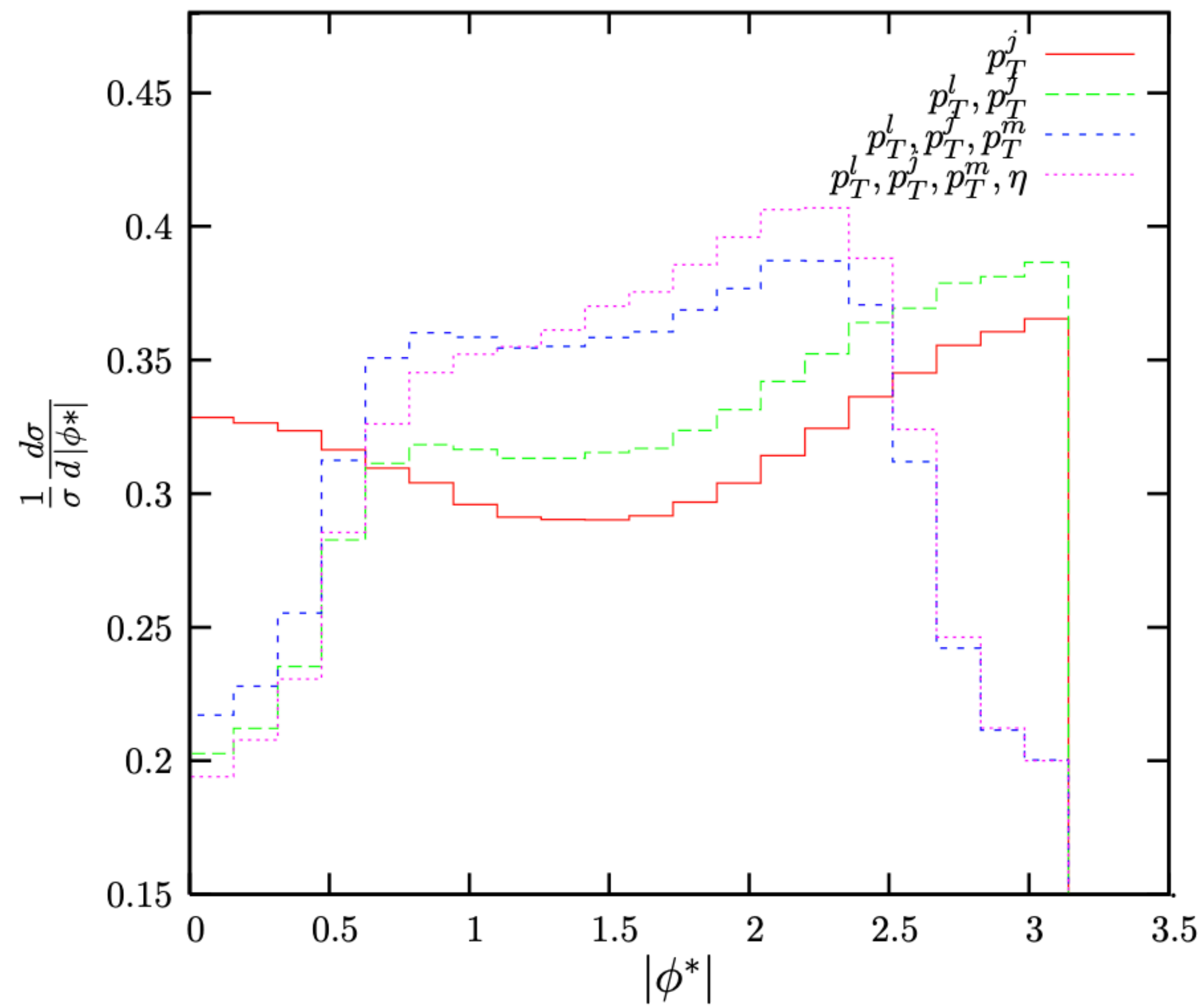


Figure 9: Normalised azimuthal angle distributions for a set of different selection cuts imposed on final-state leptons and jets for $W^+ + 1$ jet production at 7 TeV.

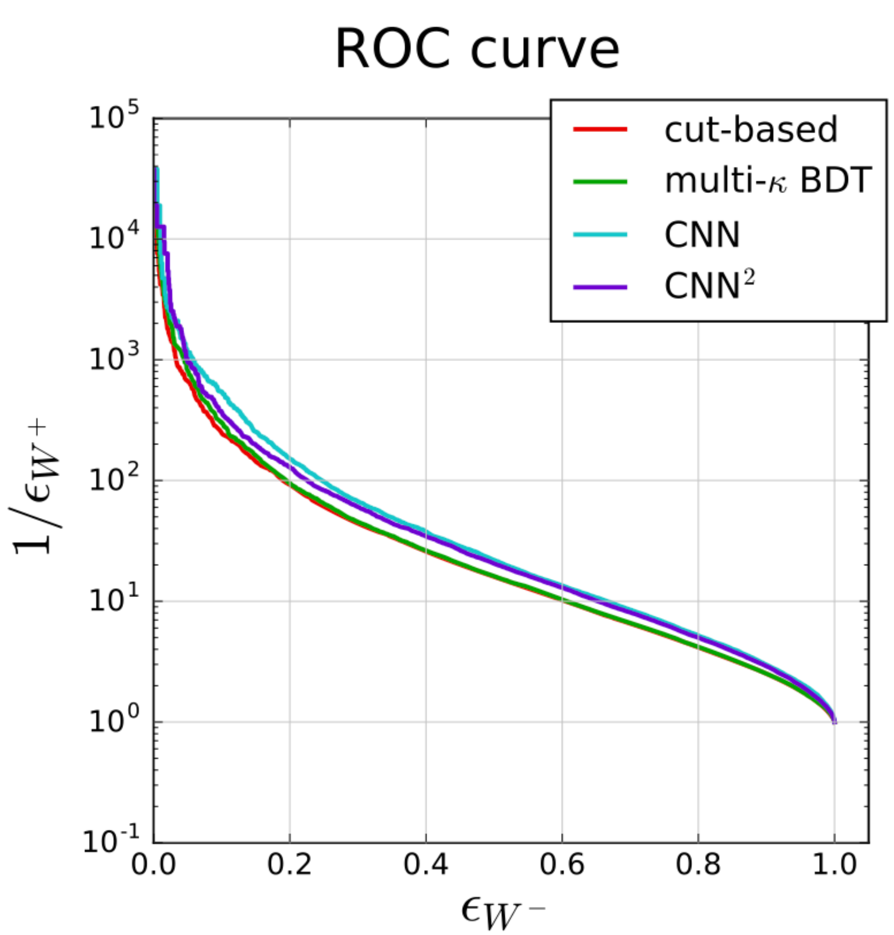
W vs. Z

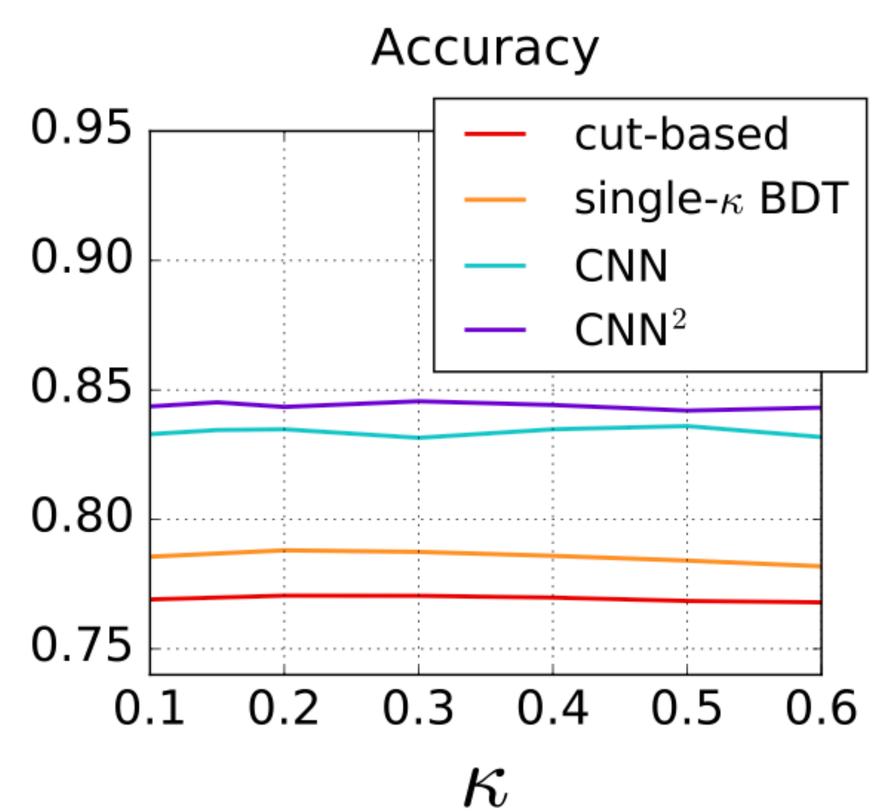
Jet charge

PhysRevD.101.053001

Additional observable : $\mathcal{Q}_\kappa = \frac{1}{(p_{T,J})^\kappa} \sum_{i \in J} q_i \times (p_T^i)^\kappa$

- Depending on κ, separation may change.
  - Need to find optimal value of κ
- Input is pT and Q_K depth=2 image





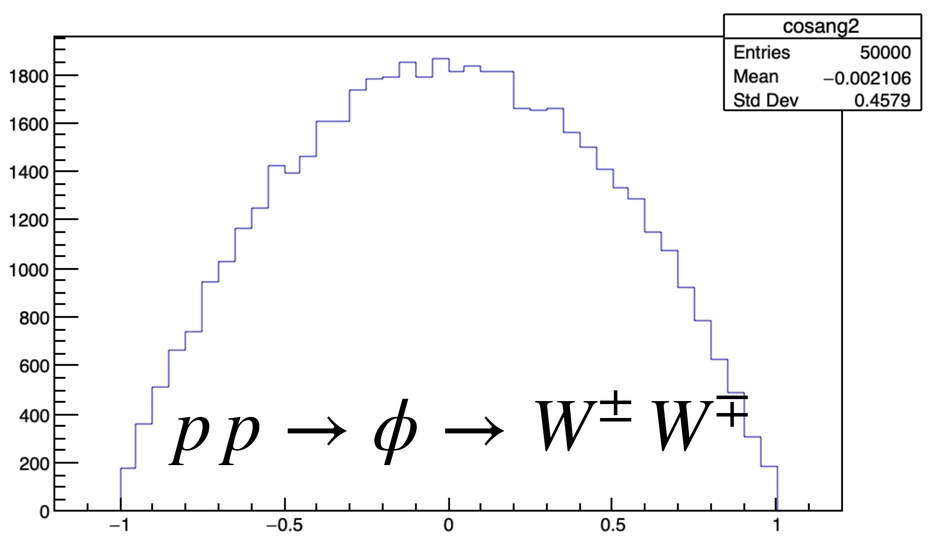
Preparing Samples

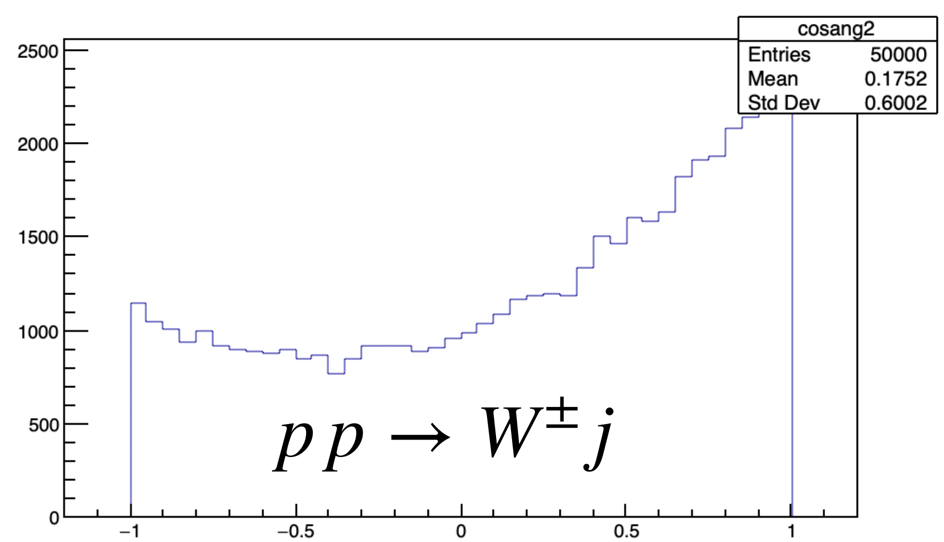
Training / Validation

Longitudinal $\quad p\,p \to \phi \to W^{\pm}\,W^{\mp}$ $\quad$ Created with heavy Higgs

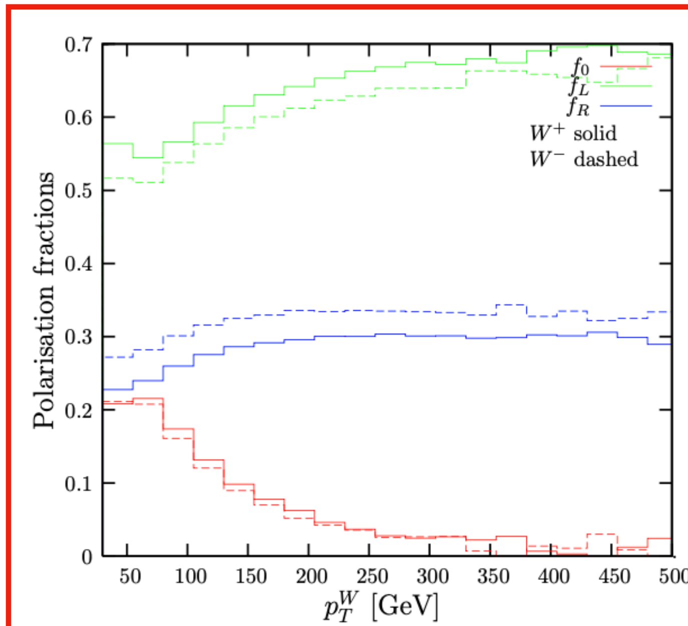
Transverse $\quad p\,p \to W^{\pm}\,j$

- **MadGraph** + **Pythia** + **Delphes**
- We separate into p_T bins of W jet: [200,300] and [400,500]
- To make sure the quality of sample, we plotted W decay in parton level





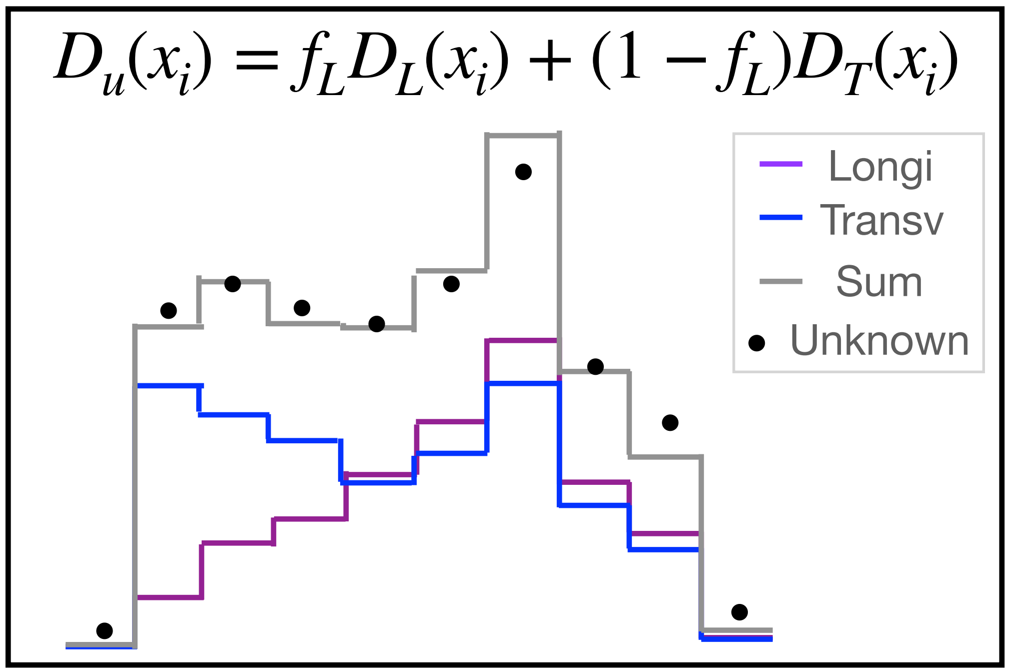
Why asymmetric?

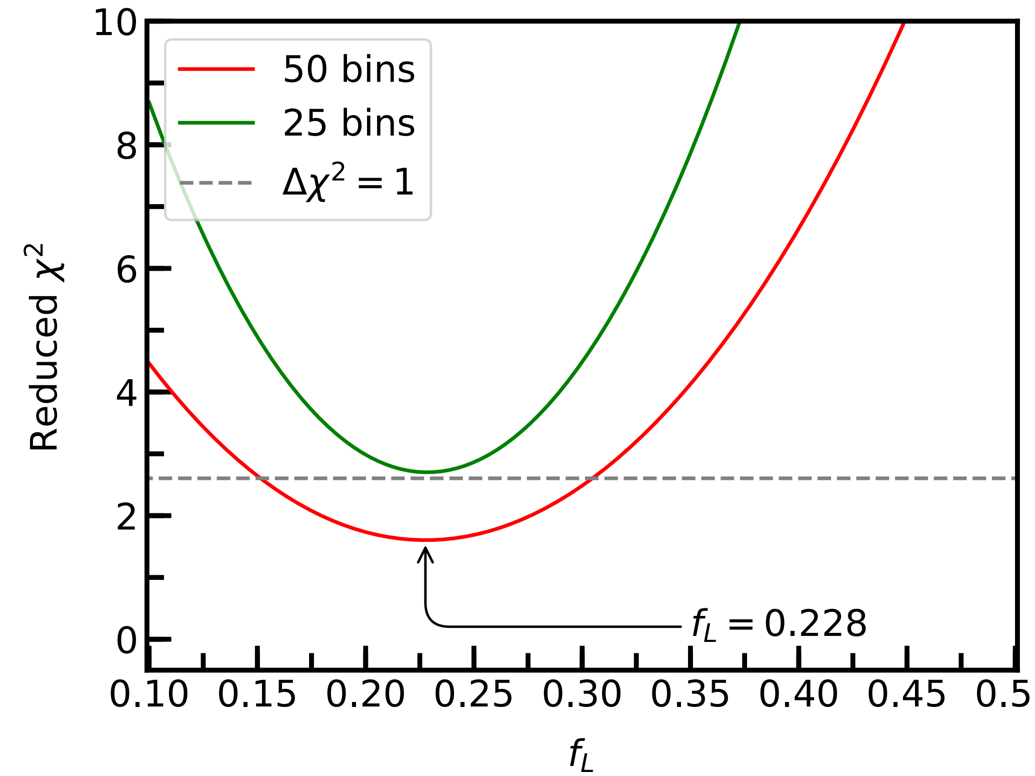


[arXiv: 1204.6427v1]

Analysis

Template fit method





- Consider each pure polarization histogram as “template” that can be applied to the unknown sample

- Fit quality is determined by χ^2 distance test

$$\chi^2(f_L) = \sum_{i=1}^{B} \frac{(O_i - N_s(f_L L_i + (1 - f_L) T_i))^2}{N_s(f_L L_i + (1 - f_L) T_i)}$$

Test on Unknown Samples

SM testing using average method

$$p\,p \to W^{\pm}Z$$

p_T range	**truth** σ_L/σ_{tot}	**predicted** f_L
[200,300]	0.265	$0.259 \pm$?
[400,500]	0.304	$0.300 \pm$?

- Output average method can predict well for both p_T bins
- Estimate error on our prediction can tell us the precision
- Truth value is calculated from **MadGraph**

Uncertainty

Small experiments

- From large test set, we randomly select subset (N number of events) to obtain f_L
- N is determined from expected number of events at particular luminosity
- At current LHC luminosity ~ 2000 events at low p_T and 200 events at high p_T
- At High Lumi LHC ~ 20k events at low p_T and 2k events at high p_T
- By iterating the process, we can obtain average value with standard deviation

	300 fb^{-1}	**3000 fb^{-1}**
[200,300]	0.044	0.010
[400,500]	0.130	0.033

Experimental Results

ATLAS result

ATLAS Result (36fb^{-1})

	f_0 Data	f_0 POWHEG+PYTHIA	f_0 MATRIX
W^+ in W^+Z	0.26 ± 0.08	0.233 ± 0.004	0.2448 ± 0.0010
W^- in W^-Z	0.32 ± 0.09	0.245 ± 0.005	0.2651 ± 0.0015
$W^\pm$ in $W^\pm Z$	0.26 ± 0.06	0.2376 ± 0.0031	0.2506 ± 0.0006
Z in W^+Z	0.27 ± 0.05	0.225 ± 0.004	0.2401 ± 0.0014
Z in W^-Z	0.21 ± 0.06	0.235 ± 0.005	0.2389 ± 0.0015
Z in $W^\pm Z$	0.24 ± 0.04	0.2294 ± 0.0033	0.2398 ± 0.0014

ATLAS Collaboration [arXiv:1902.05759]

1. Previous attempts from ATLAS collaboration to measure polarization with leptonic final states
    - Leptonic final state: small branching ratio
    - Complication in ν reconstruction
2. If we can use hadronic W, we gain more statistics but need to deal with hadronic jets