Entanglement, Bell Inequalities, and Hadronic Top Polarimetry with ML

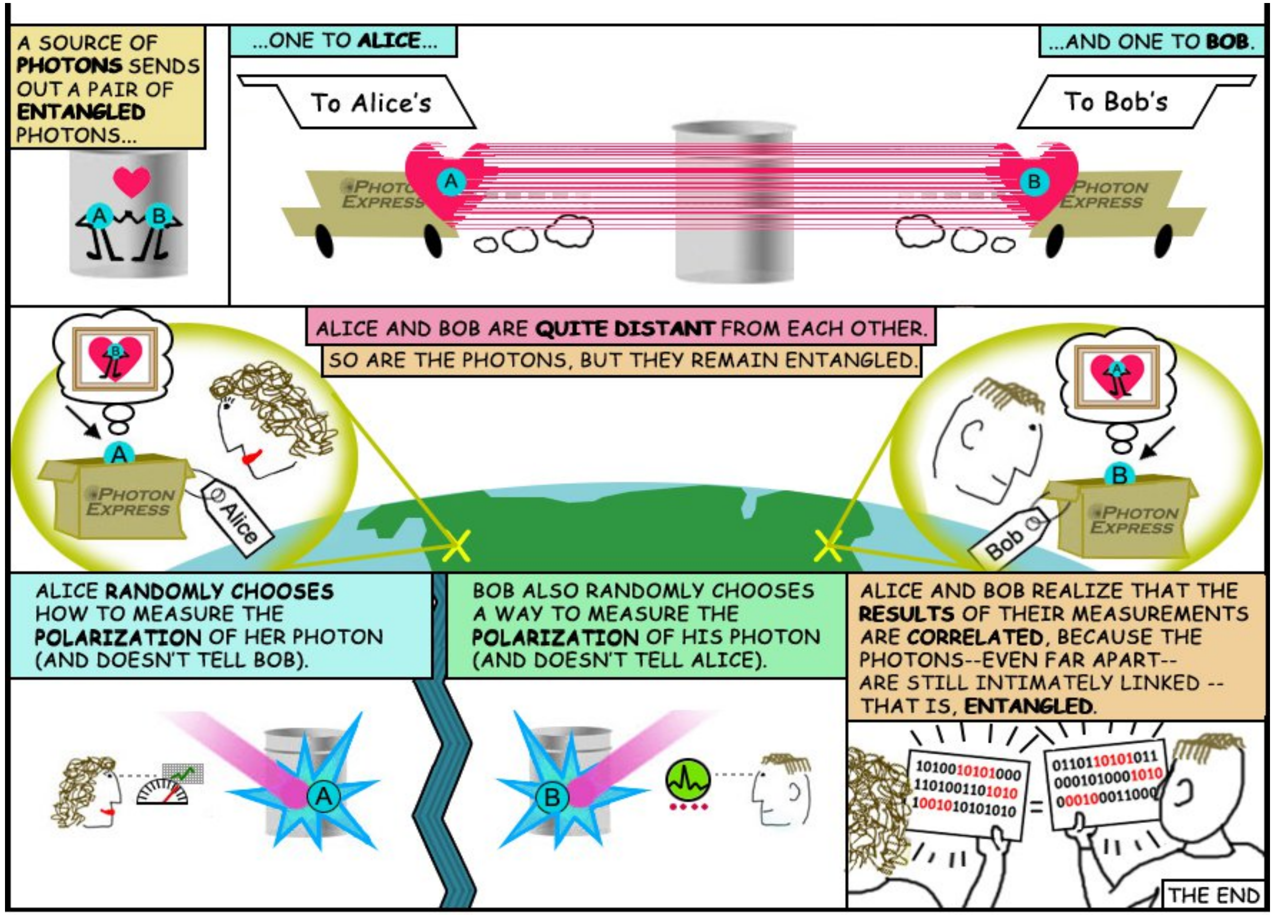
Elementary Particle Experiment Seminar, University of Washington

Alberto Navarro

Based on PRD 109, 115023,
arXiv:2407.01663,
arXiv:2407.07147
with Z. Dong, D. Gonçalves, K. Kong, and A. Larkoski

July 22, 2024

Entanglement and Bell inequality



- The presence of entanglement in itself only confirms the existence of quantum correlations that must be there because of quantum mechanics.
- Entanglement may lead to violation of the so-called Bell inequality.
- Violation of Bell inequality implies something about the physical world, i.e., non-locality.

Top and top pair as one and two-qubit system

Being a spin-1/2 particle, the top can be treat as a qubit with density matrix

$$\rho = \frac{\mathbb{1} + \sum_i B_i \sigma_i}{2}$$

top polarization vector

Easily generalized for a two-qubit system

$$\rho = \frac{\mathbb{1} \otimes \mathbb{1} + \sum_i \left(B_i \sigma_i \otimes \mathbb{1} + \overline{B}_i \mathbb{1} \otimes \sigma_i\right) + \sum_{ij} C_{ij} \sigma_i \otimes \sigma_j}{4}$$

tops polarization vectors

spin correlations

Strong top pair production satisfies P and CP

$$B_i = \overline{B}_i = 0$$

$$C_{ij} = C_{ji}$$

Our goal is to compute the correlation coefficients from data.

Entanglement

In general, for a system composed of two subsystems A and B, a quantum state is said to be separable if the mixed system is described by a density matrix

$$\rho = \sum_i p_i \rho_A^i \otimes \rho_B^i$$

An entangled state is defined as a non-separable state.

The Peres-Horodecki criterion provides a necessary and sufficient condition for entanglement. Take the transpose of the density matrix with respect to, say, subsystem B

$$\rho^{T2} = \sum_i p_i \rho_A^i \otimes \left(\rho_B^i\right)^T$$

If it results in a negative operator, the system is entangled.

For top pair production, a sufficient condition for negativity is $|C_{11} + C_{22}| - C_{33} - 1 > 0$

Afik and M. de Nova (2020)

Bell inequalities

- The Clauser, Horne, Shimony, and Holt (CHSH) inequality is a realization of the Bell-type inequalities for bipartite system

$$|\langle A_1 B_1\rangle + \langle A_2\, B_1\rangle + \langle A_1\, B_2\rangle - \langle A_2 B_2\rangle| \leq 2$$

$$\langle A_1\, B_1\rangle = Tr[\rho\,(A_1\, B_1)] = C_{ij} a_1^i b_1^j$$

With a suitable choice of vectors a_1, b_1, a_2, b_2 the violation of the CHSH inequality requires

$$|C_{ii} \pm C_{jj}| > \sqrt{2}$$

Aguilar and Casas (2022)

Why top quarks?

The top quark is special

$$\tau_{top} \approx 5 \times 10^{-25} s$$

$$\tau_{had} \approx 2 \times 10^{-24}\ s$$

$$\tau_{flip} \approx 10^{-21}\ s$$

Can we use the properties of the top quark to explore entanglement and violation of Bell inequalities at the LHC? YES!

Top polarization

The direction of the top quark decay products correlate with its polarization axis as

spin analyzing power

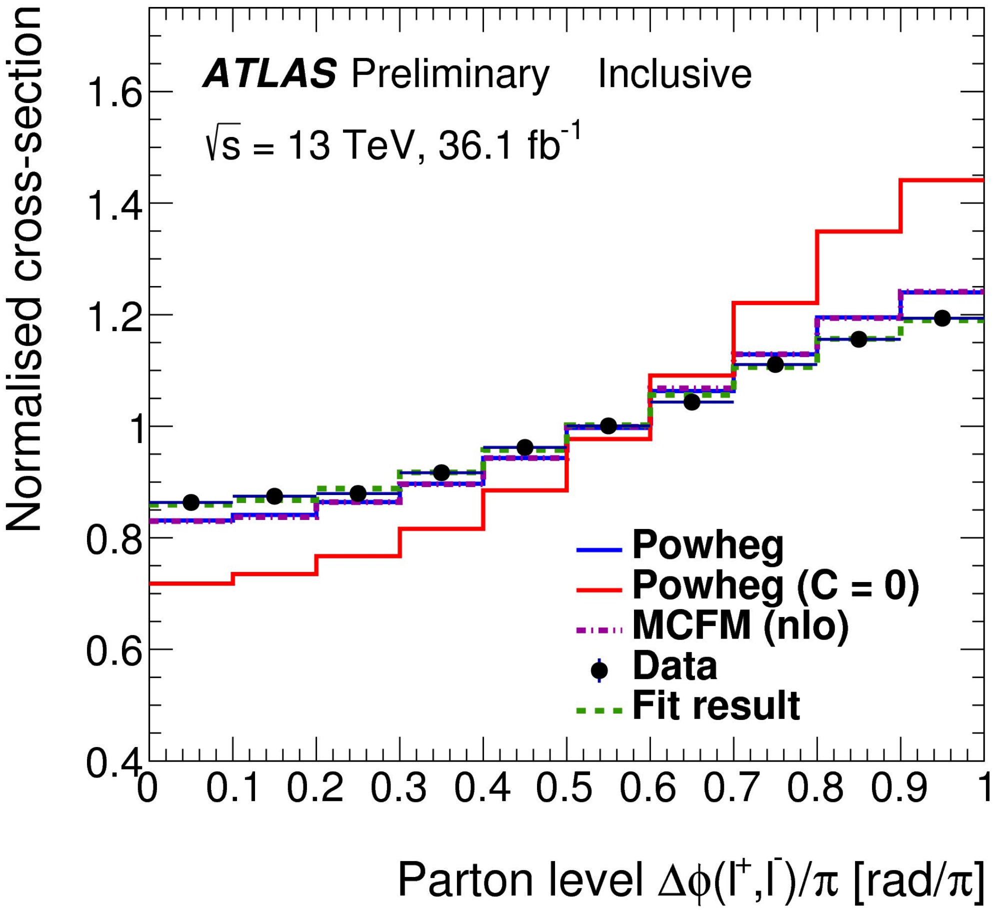
$$\frac{1}{\Gamma}\frac{d\Gamma}{d\cos\xi_k} = \frac{1}{2}(1 + p\beta_k \cos\xi_k) \qquad \beta_k = \begin{cases} +1 \ for \ \ell^+ \ or \ \bar{d}-quark \\ -0.31 \ for \ \bar{\nu} \ or \ u-quark \\ -0.41 \ for \ b-quark \end{cases}$$

polarization of the ensemble

- We can obtain the top spin information from angular distributions involving its decay products.
- Charged leptons and down-type quark directions are the best proxies for the top spin.

Spin correlations in top-pair production

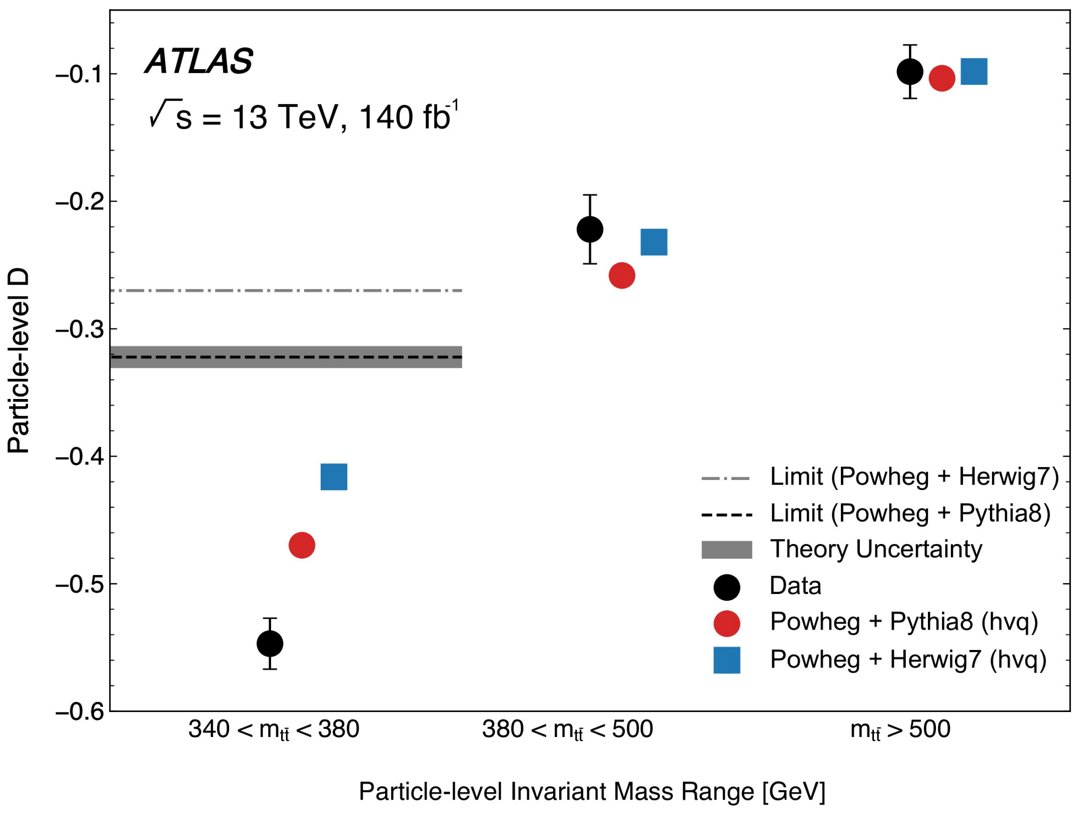
- Top quark polarization allows us to have a look at spin correlations in top-pair production.
- Spin correlations in top-pair production have been known and measured for a long time.
- Entanglement and violation of the Bell inequalities can be translated into extra conditions on the spin correlations.



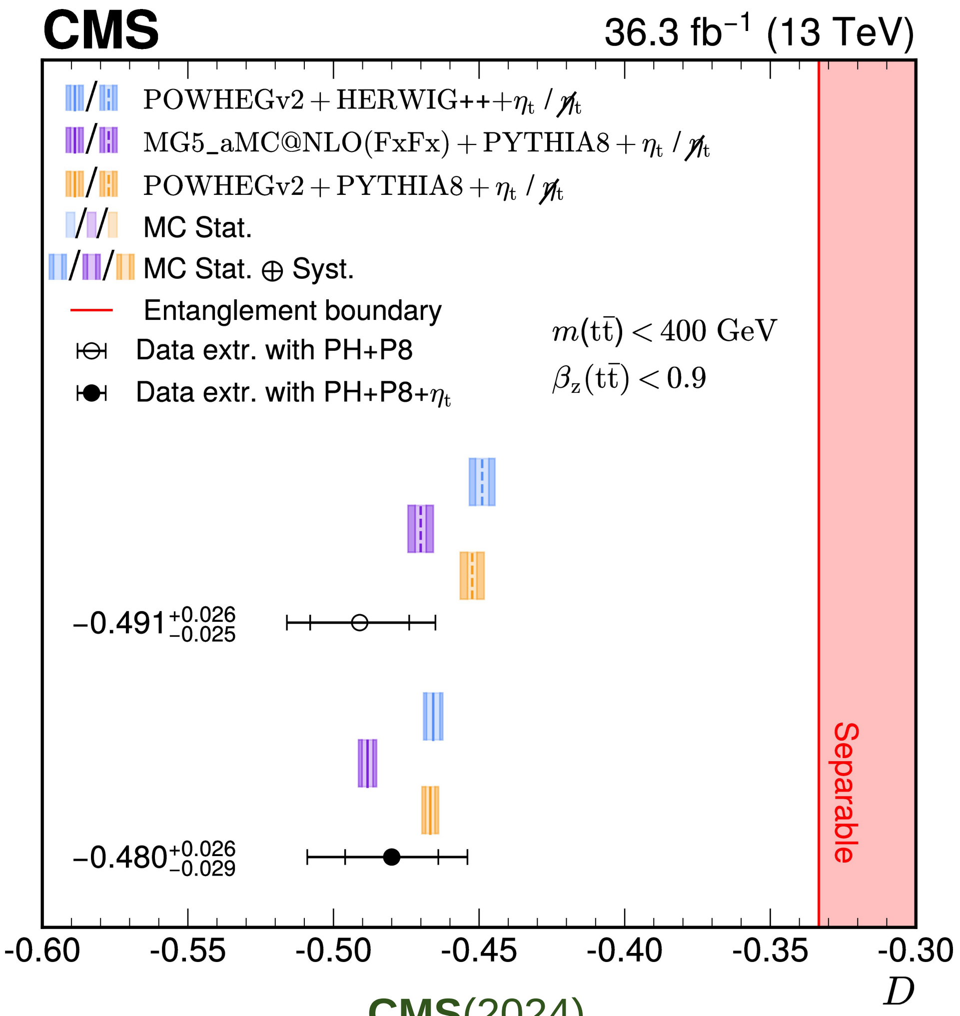
ATLAS (2020)

Measurement of entanglement in di-lepton final state

Both ATLAS and CMS have measured entanglement at threshold in di-lepton final state.



ATLAS (2023)



CMS(2024)

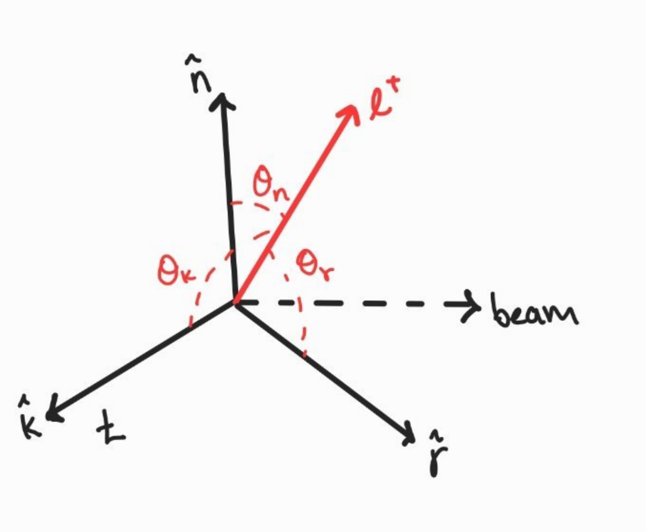
Density matrix reconstruction

The elements of the correlation matrix are extracted from forward-backward asymmetries

$$C_{ij} = \frac{4}{\beta_a \beta_b} \frac{N\left(\cos\theta_a^i \ \cos\overline{\theta}_b^j > 0\right) - N\left(\cos\theta_a^i \ \cos\overline{\theta}_b^j < 0\right)}{N\left(\cos\theta_a^i \ \cos\overline{\theta}_b^j > 0\right) + N\left(\cos\theta_a^i \ \cos\overline{\theta}_b^j < 0\right)}$$

The axes are chosen in the so-called helicity basis

$$\hat{k} = top\ direction,\ \hat{r} = sign(\cos\theta)\frac{\hat{p} - \hat{k}\cos\theta}{\sin\theta},\ \hat{n} = \hat{k}\times\hat{r}$$

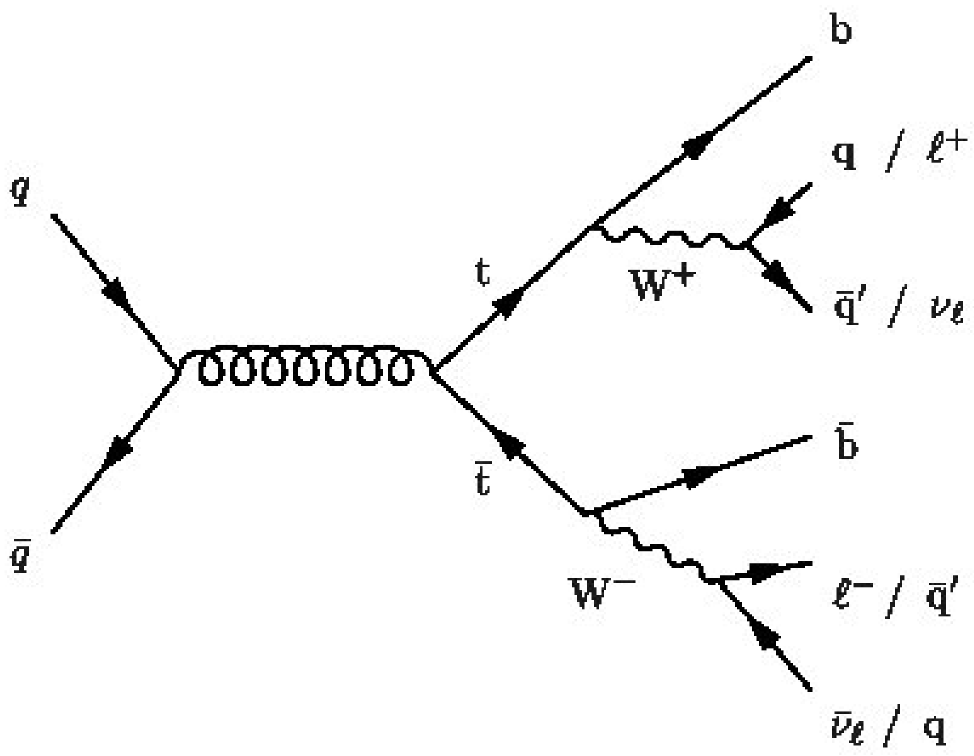


Charged lepton used as proxy for leptonic top spin. What about the hadronic side?

Ups and downs of semileptonic tops

Z. Dong, D.Gonçalves, K. Kong, AN (2023)
see also Han, Low and Wu (2023)

- Semileptonic channel offers larger statistics.
- Hadronic top reconstruction can be addressed using jet substructure techniques.
- The down quark is the best polarimeter for the hadronic top, but tagging it in a collider environment is challenging.

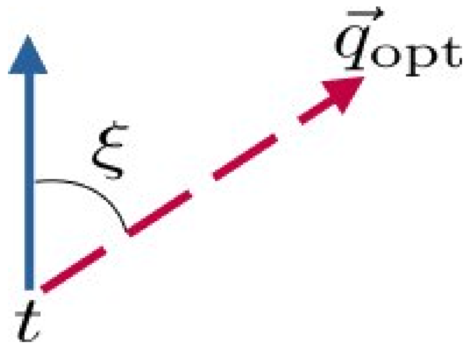


Hadronic top polarimeters

- A possible choice is the softest jet in the top rest frame. This gives a spin analyzing power equals to 0.5.

- It is possible to define a optimal polarimeter by using all the kinematic information from the top decay.

- This optimal polarimeter gives a spin analyzing power equals to 0.64. Tweedie (2014)



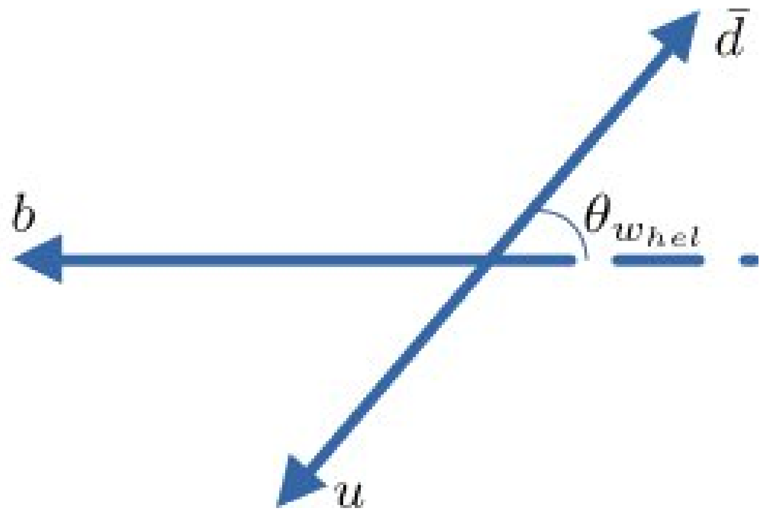
It is possible to build a better polarimeter using information from jet constituents. Will show how in the second part of the talk.

Optimal hadronic direction

In the top rest frame, the soft and hard-quark each have a probability of being the d-quark

$$\text{prob}\left(d \to q_{\text{hard}}\right) = \frac{\rho\left(|c_{w_{\text{hel}}}|\right)}{\rho\left(|c_{w_{\text{hel}}}|\right) + \rho\left(-|c_{w_{\text{hel}}}|\right)}$$

$$\text{prob}\left(d \to q_{\text{soft}}\right) = \frac{\rho\left(-|c_{w_{\text{hel}}}|\right)}{\rho\left(|c_{w_{\text{hel}}}|\right) + \rho\left(-|c_{w_{\text{hel}}}|\right)}$$

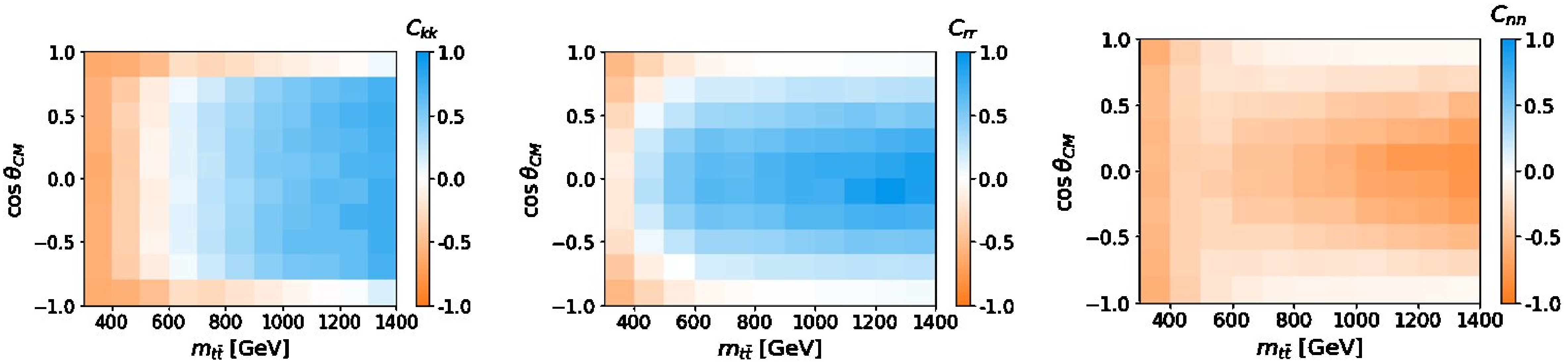


The optimal hadronic direction is defined as the weighted sum of the soft and hard-quark directions

Tweedie (2014)

$$\vec{q}_{\text{opt}} = \text{prob}\left(d \to q_{\text{hard}}\right)\hat{q}_{\text{hard}} + \text{prob}\left(d \to q_{\text{soft}}\right)\hat{q}_{\text{soft}}$$

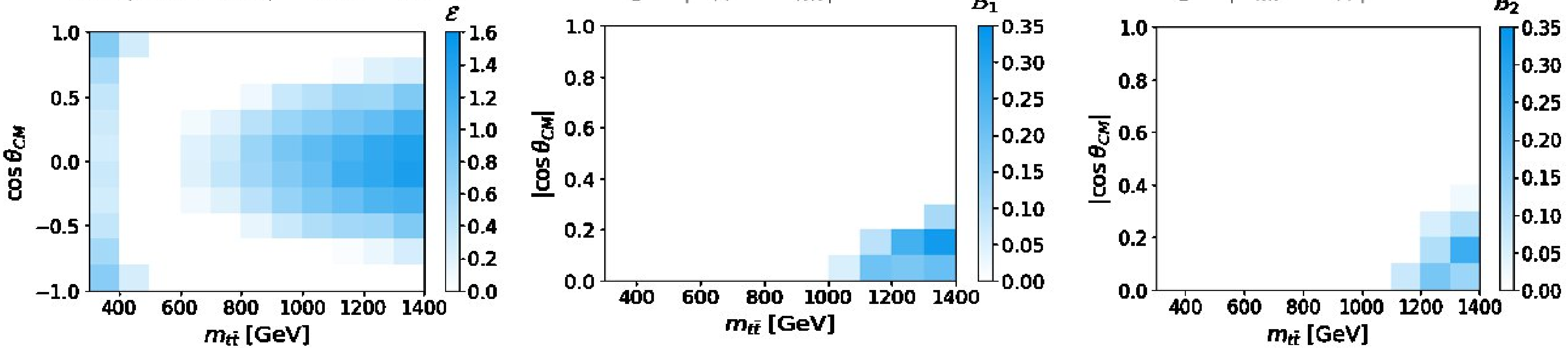
Parton-level distributions



$$\mathcal{E} \equiv |C_{kk} + C_{rr}| - C_{nn} - 1 > 0$$

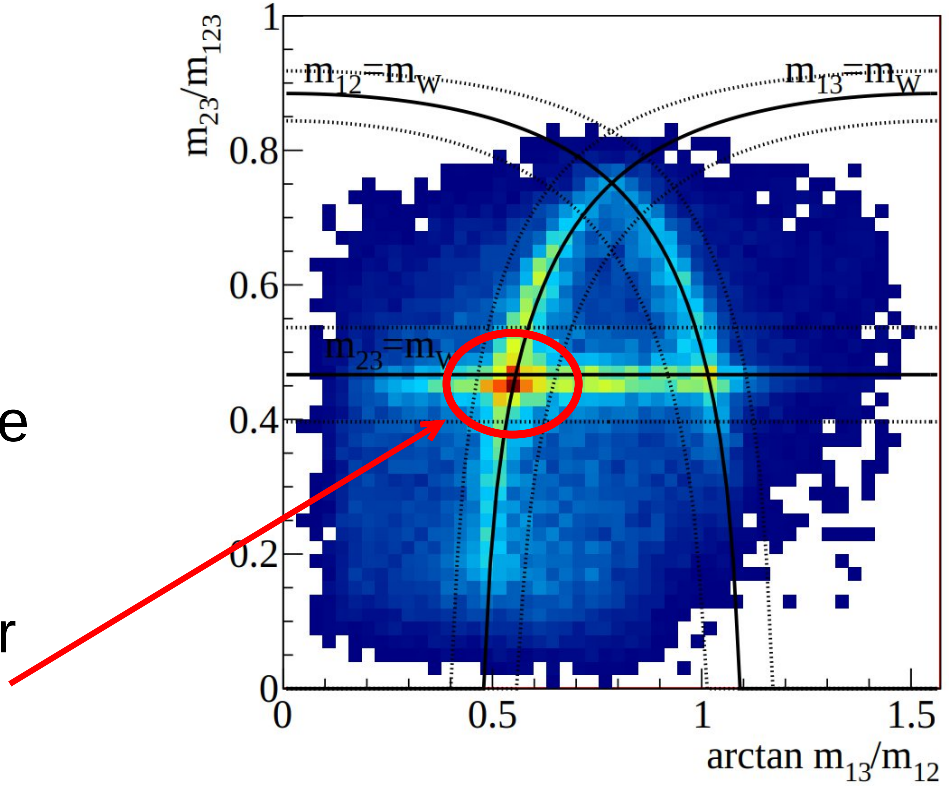
$$\mathcal{B}_1 \equiv |C_{rr} - C_{nn}| - \sqrt{2} > 0$$

$$\mathcal{B}_2 \equiv |C_{kk} + C_{rr}| - \sqrt{2} > 0$$



Hadronic top reconstruction

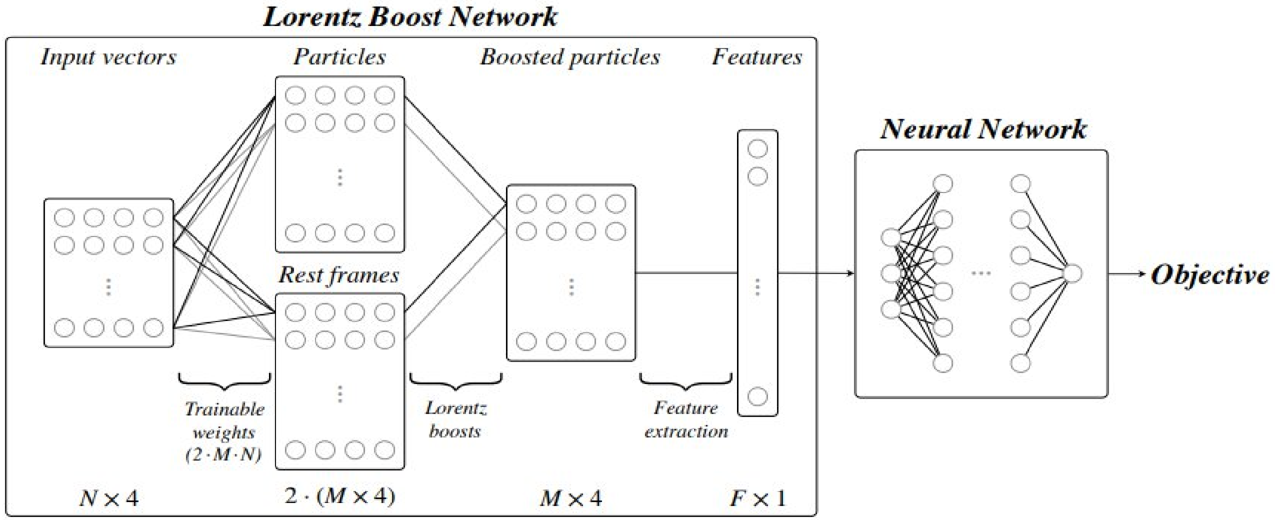
- Reconstruct fatjets with CA algorithm and R=1.5, requiring at least one with pT>150 GeV and |η|<3.
- We require that one of the fat jets to be tagged with the HEPTopTagger.
- HepTopTagger returns 3 sub-jets with a b-candidate based on W-mass criteria.
- Some ambiguity on selecting the two jets that better resemble the W.

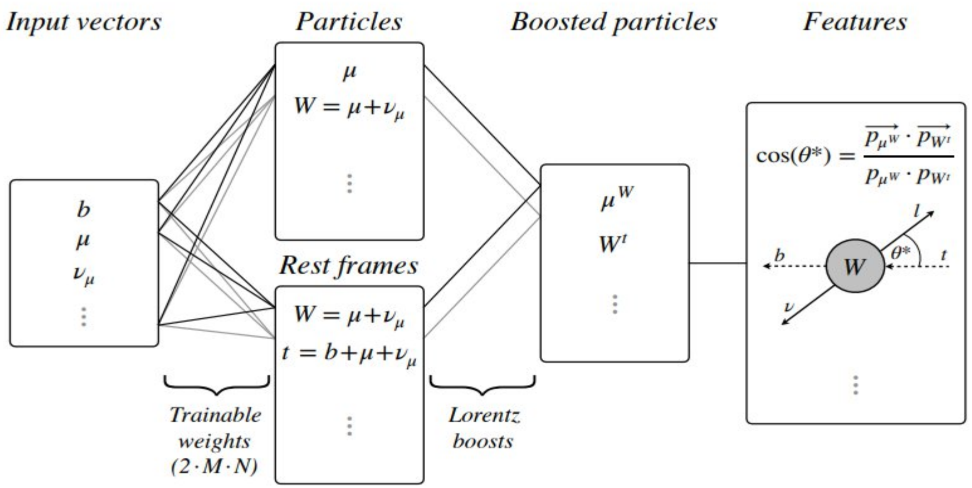


Plehn, Spannowsky, Takeuchi, Zerwas (2010)

Identifying b-jet with ML

We use the Lorentz Boost Network for identifying the b-jet.





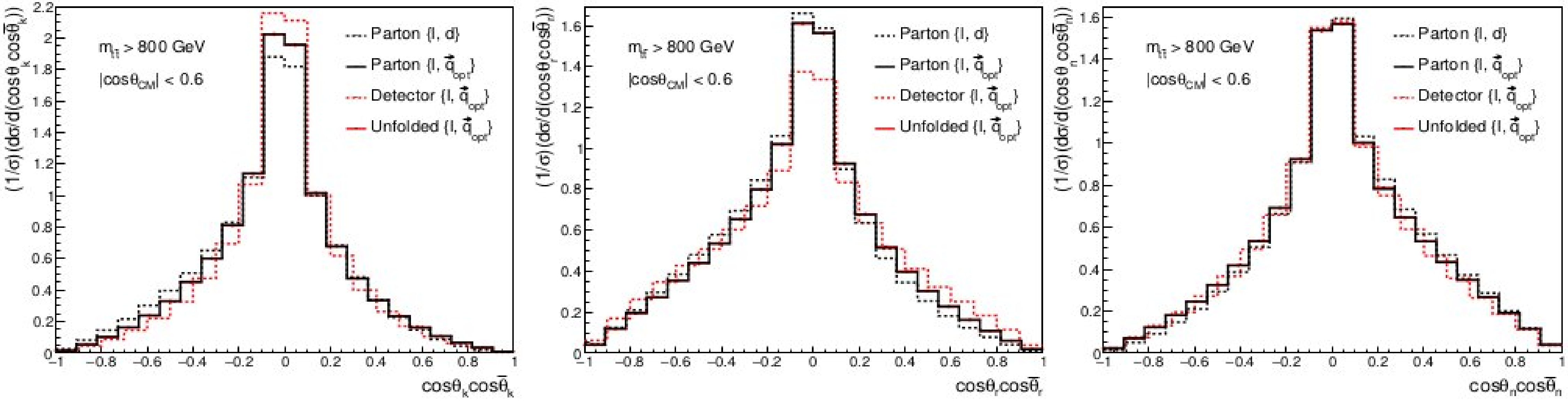
Erdmann, Geiser, Rath, Rieger (2018)

- Input 4-momenta of sub-jets (pt-ordered) and outputs categorical labels.
- Number of hypothetical particles or rest frames set to 12.
- Energy, pT, η, ϕ, mass and angles between each pair of jets are further fed into a simple DNN.

Neutrino reconstruction with LBN

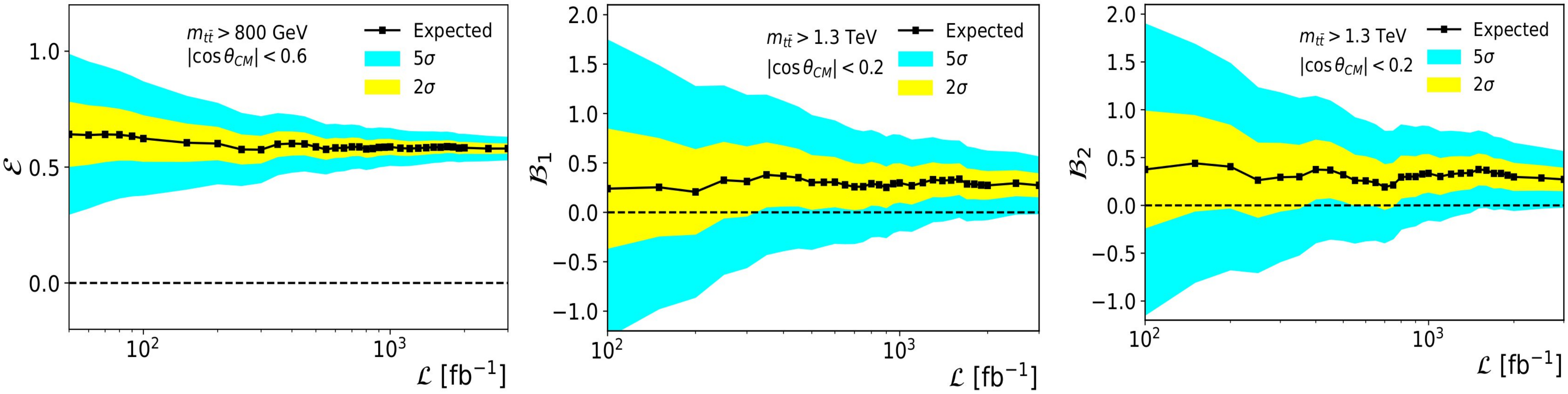
- The momenta of b-jet, lepton and neutrino (x,y=missing energy and E=pz=0) are used and input features.
- Output layer as hadronic case, i.e., three numbers corresponding to the x, y, and z components of neutrino momentum.

$$L = \text{Mean squared loss} + \lambda_1 (m_t - m_{b\ell\nu})^2 + \lambda_2 (m_W - m_{\ell\nu})^2$$



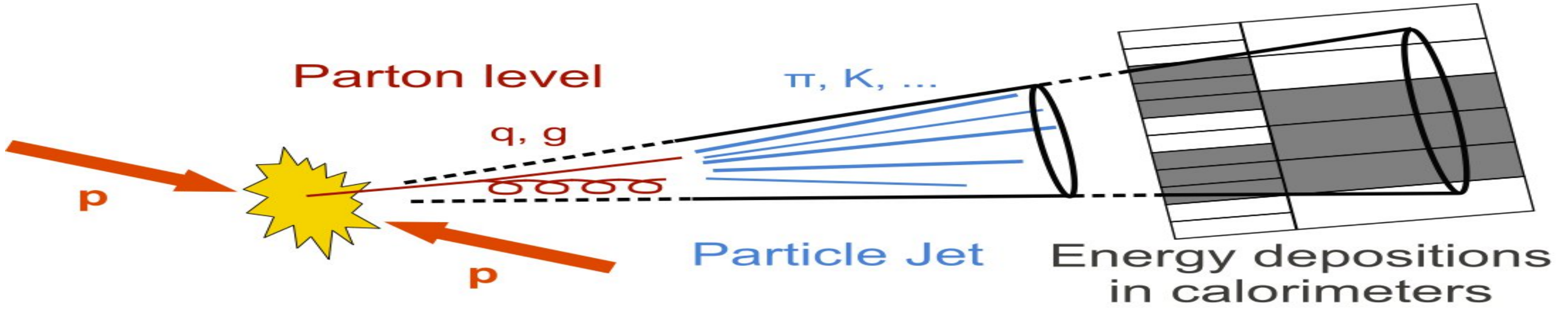
Sensitivity at the HL-LHC

The reconstructed distributions are then unfolded with TSVDUnfold.



Possible path for improvement

- Measuring the hadronic top polarization is possibly the biggest challenge.
- The optimal direction does a good job, but it uses only kinematic information of the jets.
- At particle-level, jets contain more information than just momentum.

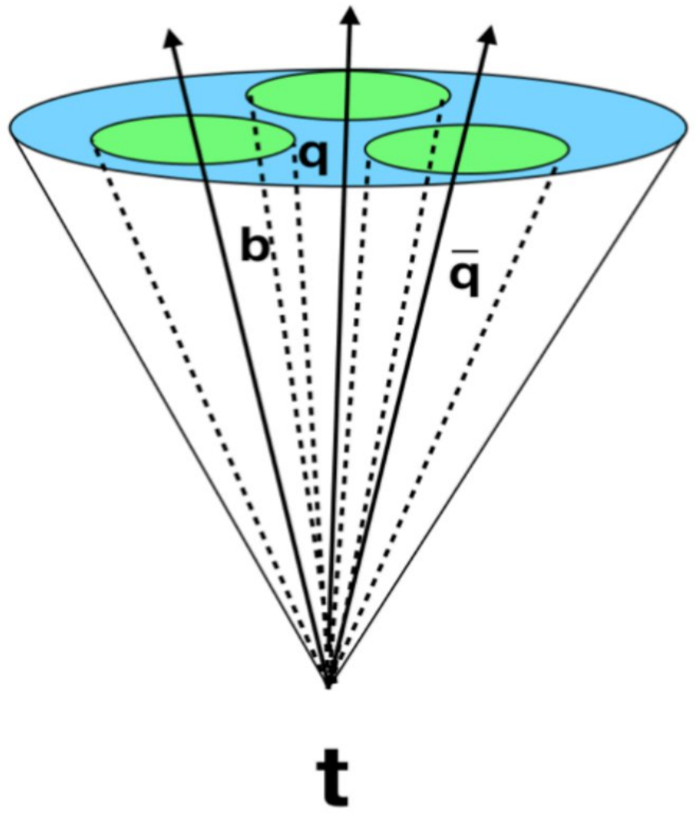


- Kinematics, charge and PID (to some extend) of the jet constituents can also be measured. How can all of this be incorporated when computing the optimal hadronic direction?

Z. Dong, D. Gonçalves, K. Kong, A. Larkoski and AN (2024)

ML approach

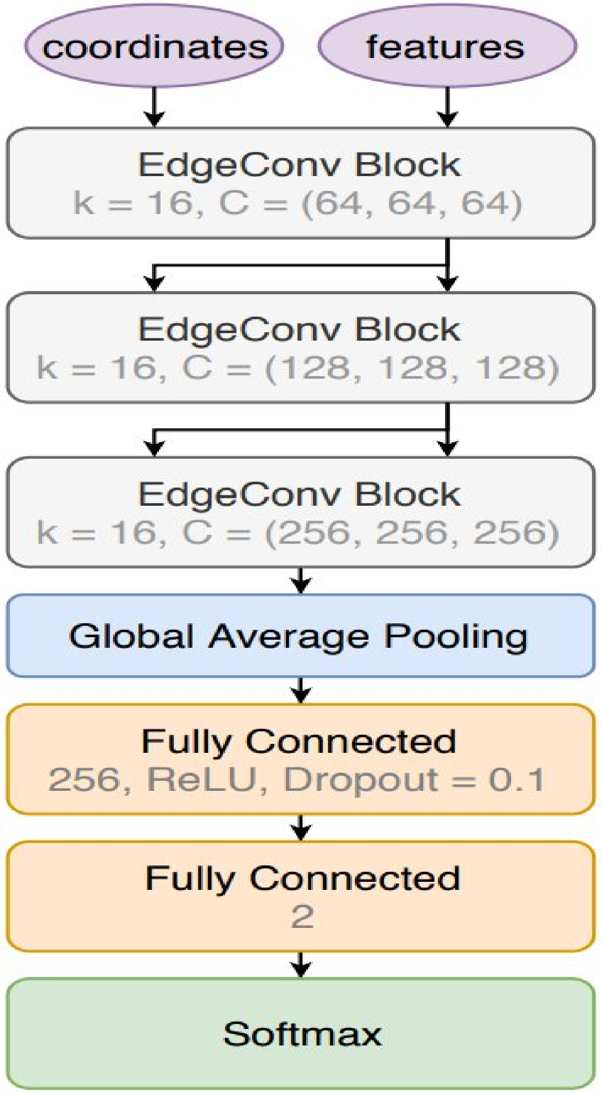
- Light jet identification within the top jet.
- Input the jet constituent momenta and charge information for each of the subjets.
- Train the neural network to identify the down-type jet.
- Interpret the neural network score as the probability of each jet being down-type.



Z. Dong, D. Gonçalves, K. Kong, A. Larkoski and AN (2024)

ParticleNet

- Idea is to represent a jet as an unordered, permutation invariant set of particles (a particle cloud).
- Analogous to the point cloud representation of 3D shapes used in computer vision.
- Uses the edge convolution operation which can be viewed as a mapping from a point cloud to another point cloud with the same number of points.
- Has been used for top and quark-gluon tagging.



Qu and Gouskos (2019)

GNN structure and input features

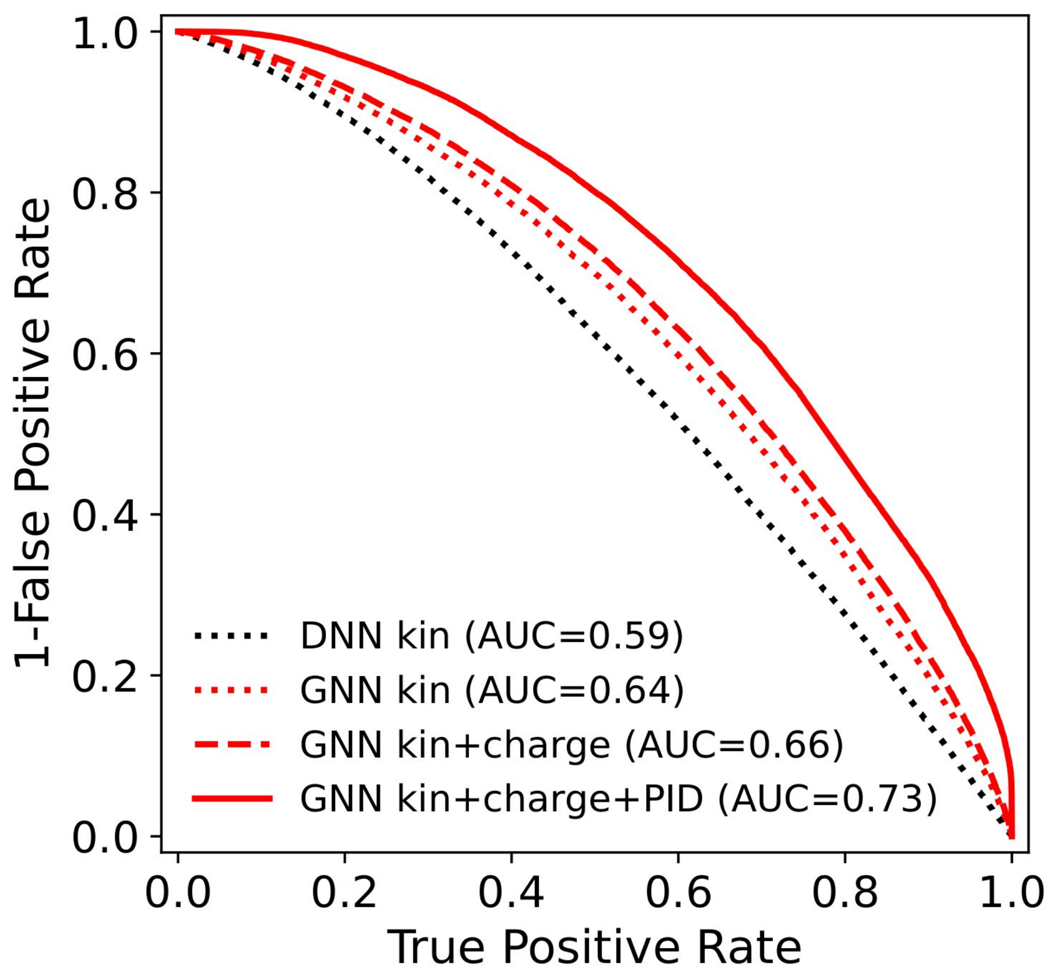
Variable	Definition
$\Delta\eta_t$	difference in pseudorapidity between the particle and the top jet axis
$\Delta\phi_t$	difference in azimuthal angle between the particle and the top jet axis
$\Delta\eta_j$	difference in pseudorapidity between the particle and the subjet axis
$\Delta\phi_j$	difference in azimuthal angle between the particle and the subjet axis
$\log p_T$	logarithm of the particle's p_T
$\log E$	logarithm of the particle's Energy
q	electric charge of the particle
isElectron	if the particle is an electron
isMuon	if the particle is a muon
isPhoton	if the particle is a photon
isChargedHadron	if the particle is a charged hadron
isNeutralHadron	if the particle is a neutral hadron

- We modified based on the ParticleNet architecture by utilizing three separate graph convolutions instead of one, corresponding to each of the jet inputs.
- The three graphs are then pooled and concatenated.
- Additional features for the overall top jet can also be fed into the linear layers.

Performance of the network

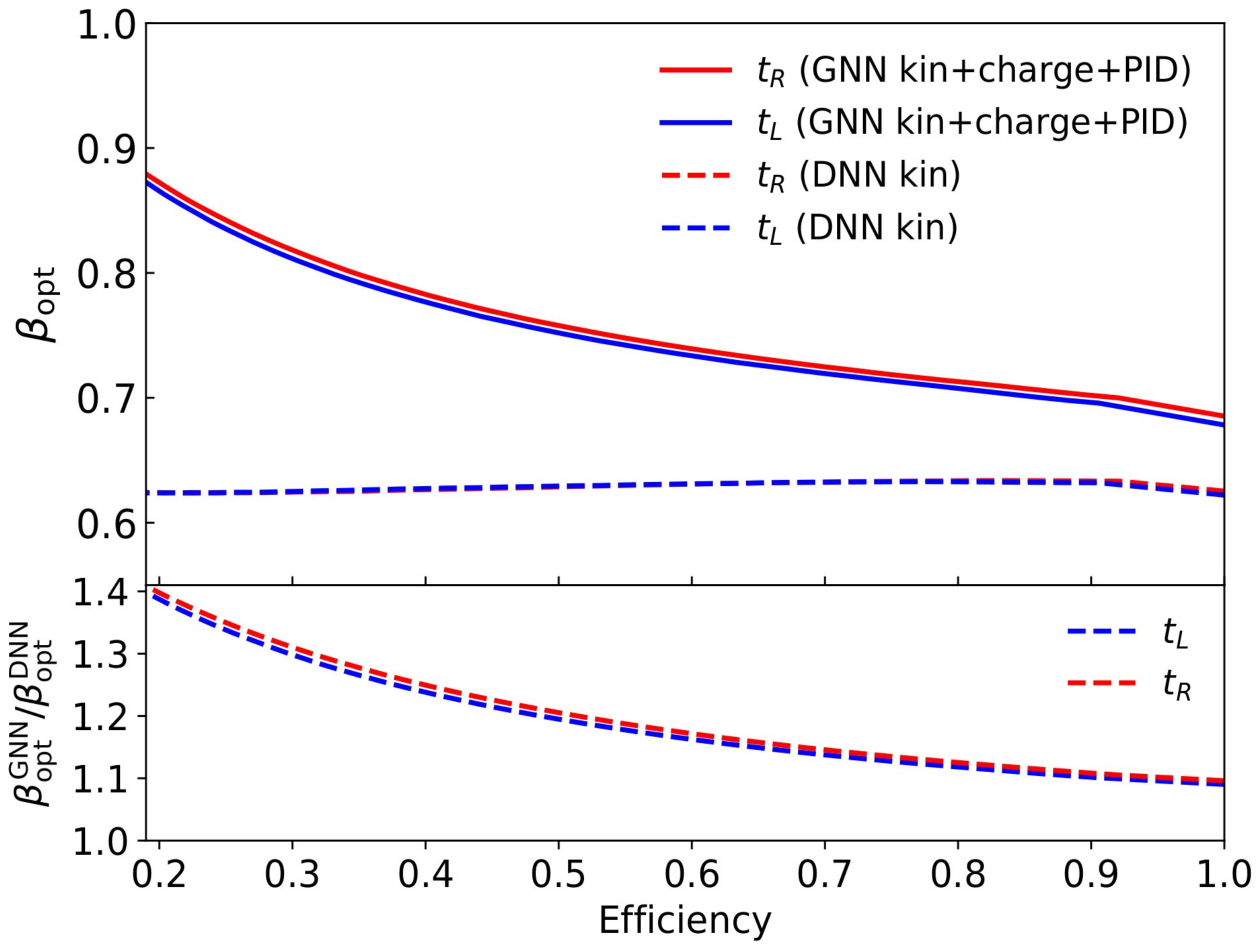
- The ROC curve of the network trained and tested on the unpolarized top data.

- “Kinematics” curve is the baseline constructed using a fully connected DNN with jet momenta input.



Optimal hadronic direction with ML

- The new optimal direction is constructed using the network score as the probability of the down-quark being the soft jet.
- We can apply cuts on the network score to obtain an efficiency and compute the spin analyzing.
- Results improve ~ 40%!



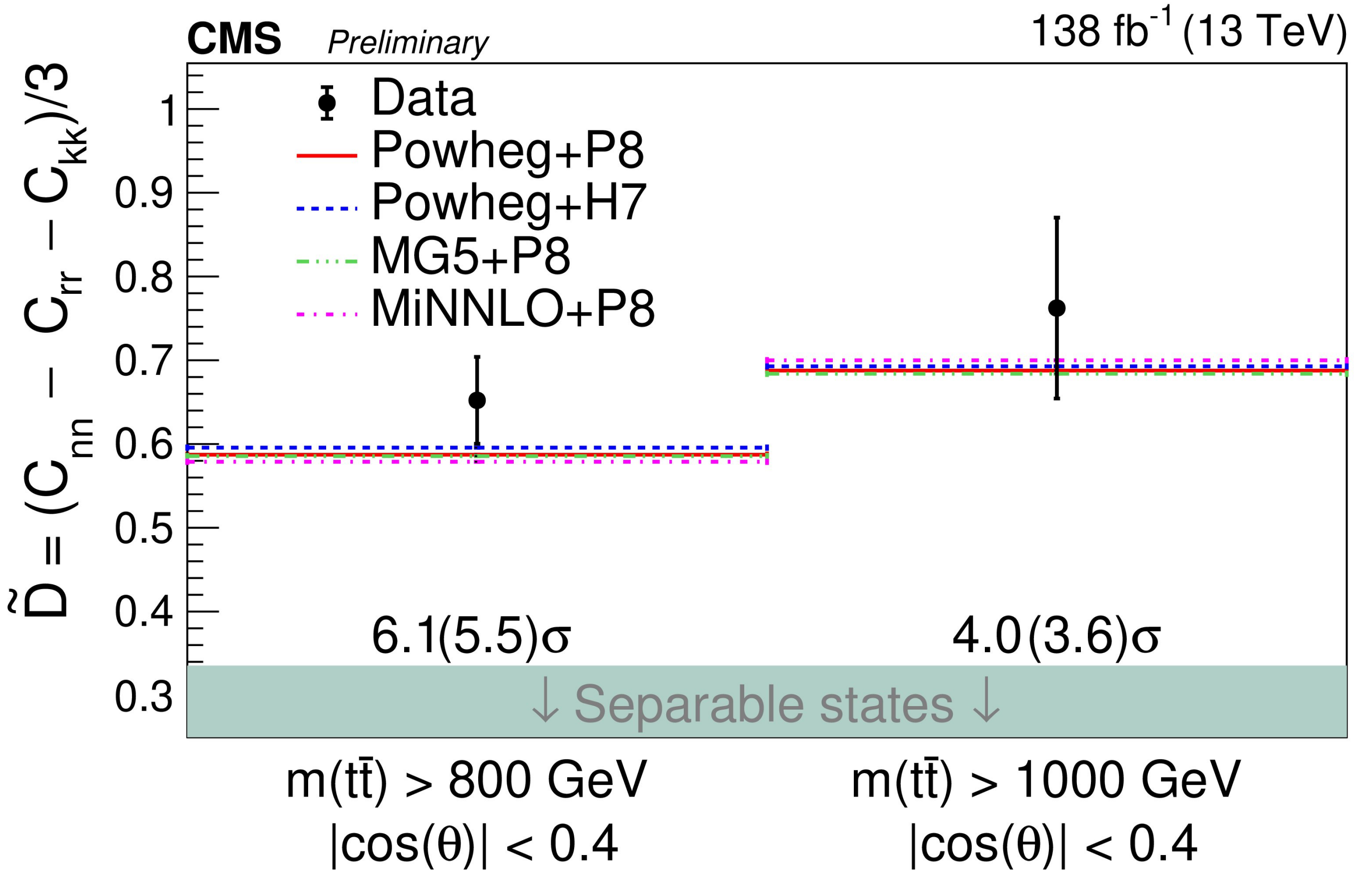
Summary

- Entanglement is present at threshold and high invariant mass regions. The violation of the Bell’s inequality may be addressed in the high invariant mass region where semi-leptonic channel may play an important role.
- The optimal hadronic direction allows us to recover the hadronic top spin quite well.
- Hadronic top polarization can be improved by incorporating information of jet constituents with the help of GNN.
- GNN also provides a way to make selections on the events to improve the spin analyzing power, which means better top spin measurements.

Backup slides

Measurement of entanglement in semi-leptonic channel

Recent measurement in semi-leptonic channel with boosted tops by CMS.



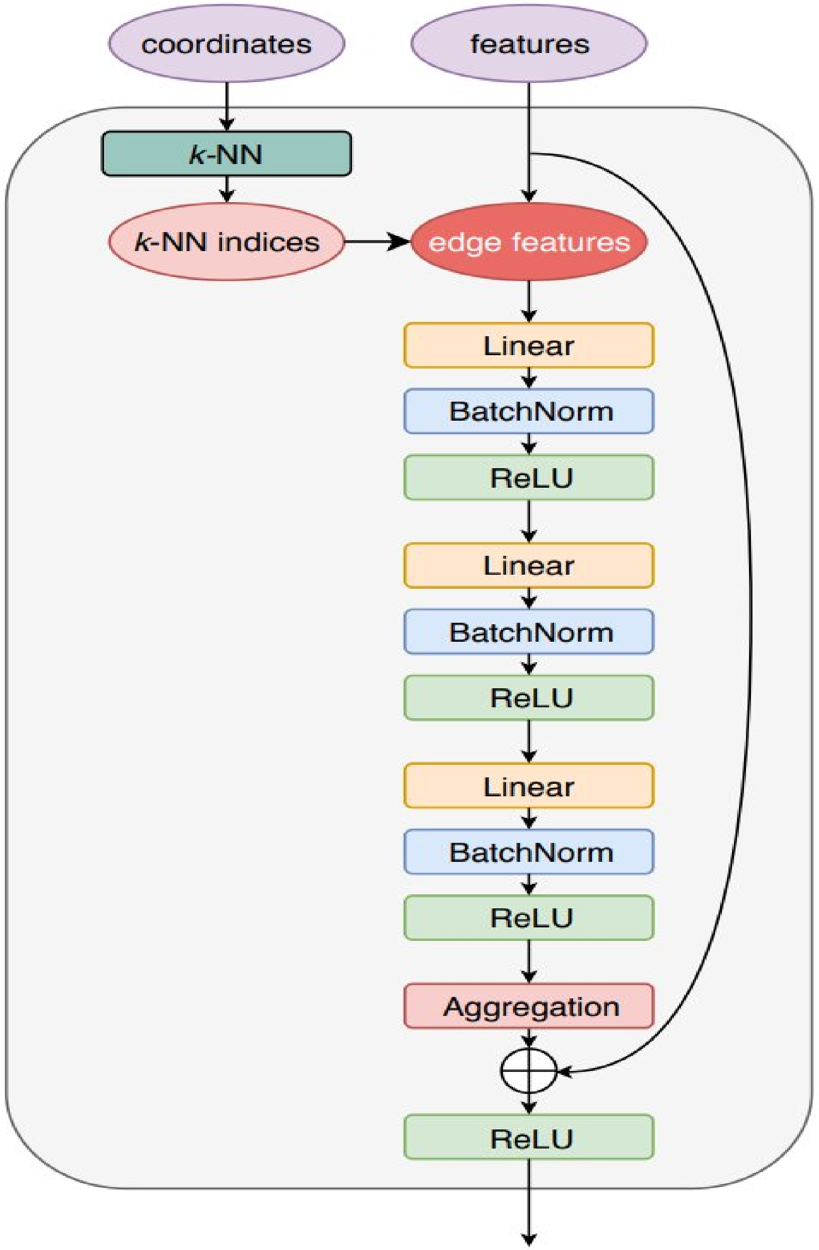
Leptonic top reconstruction

- Remaining hadronic activity is clustered with anti-kt algorithm and R=0.4 with pT>30 GeV and $|\eta|<3$. Demand a b-tagged jet.
- Neutrino momentum reconstruction can be done by using on-shell condition of W or top to find z-component. Sometimes leads to complex solutions.
- Another alternative is to find the three components of momentum via

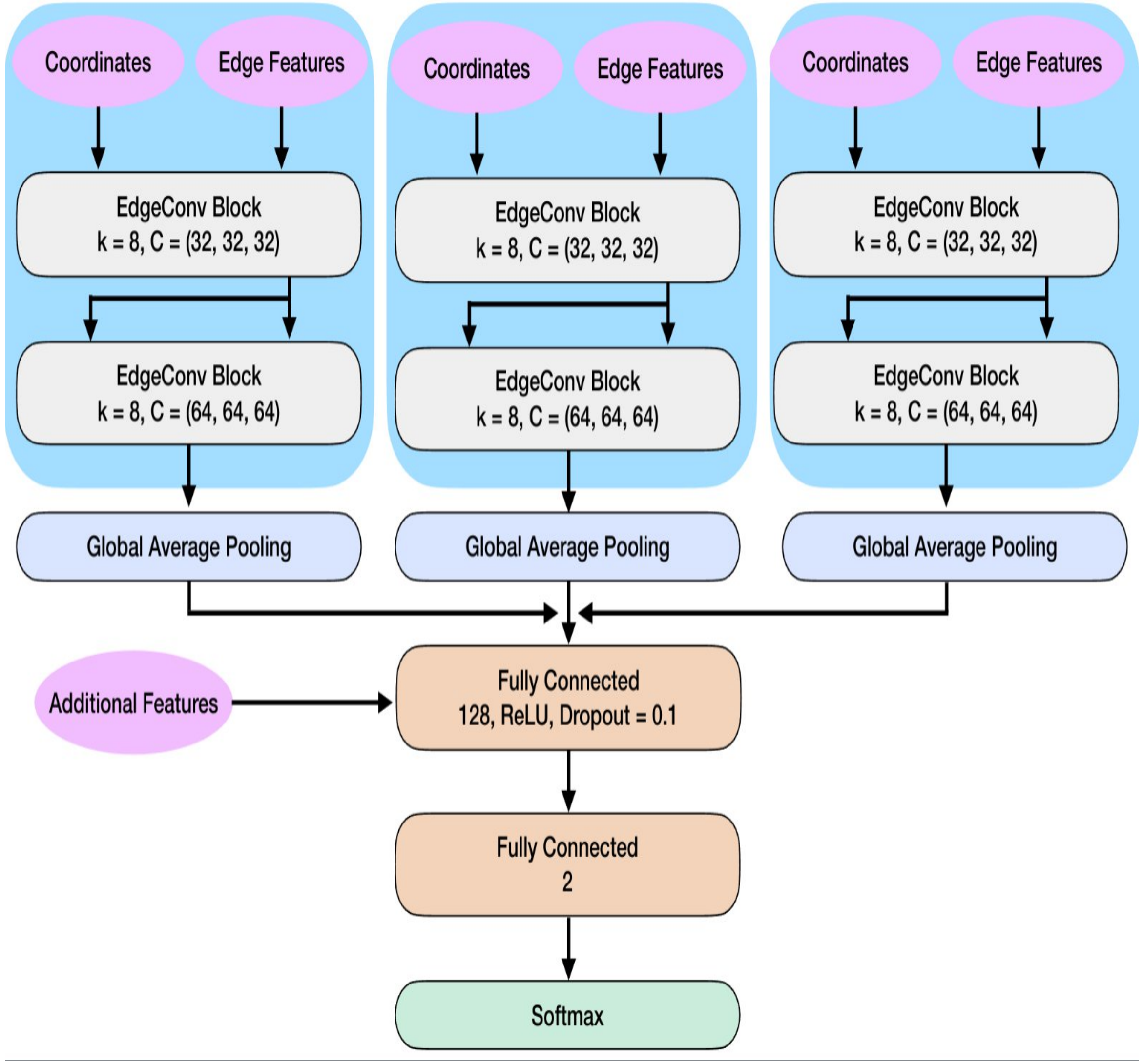
$$\chi^2 = \frac{(m_t - m_{b\ell\nu})^2}{\sigma_t^2} + \frac{(m_W - m_{\ell\nu})^2}{\sigma_W^2} + \frac{(\vec{\not{p}}_T - \vec{p}_{\nu T})^2}{\sigma_{MET}^2}$$

- Results can significantly differ from parton-level angular distributions.
- We found LBN does a better job.

EdConv block



GNN structure



Data preparation for GNN

- We generate 14TeV $pp \to t\bar{t} \to \ell^{\pm}\nu 2b\, 2j$ events using MG5, with no cuts except for pTt > 200 GeV.
- Three sets of samples where the top quark is unpolarized, left hand polarized and right hand polarized in the tt гъrest frame.
- Parton shower and hadronization are done with PYTHIA8 without MPI.
- Identify the top jet using CA algorithm with R = 1.5, and pT > 250 GeV. And decluster following to find the subjets.
- We match the hadron level jets with true parton level momenta, by using the smallest ΔR between the two.