paper link

journal club
06 Dec 2024

Hadronic Top quark polarimetry with particleNet

"playing tag with the down quark at the ttbar playground"

Jules Vandenbroeck

GHENT UNIVERSITY

CMS

Motivation

$$\frac{1}{\Gamma}\frac{d\Gamma}{d\cos\theta_i} = \frac{1}{2}(1 + \beta_i p \cos\theta_i)$$

Motivation

partial decay width of top quark

angular distribution
of the decay products

$$\frac{1}{\Gamma}\frac{d\Gamma}{d\cos\theta_i} = \frac{1}{2}(1 + \beta_i p \cos\theta_i)$$

angle of the top quark decay
product with the top spin axis in
the top quark rest frame

Motivation

partial decay width of top quark

average top quark polarization

angular distribution of the decay products

$$\frac{1}{\Gamma}\frac{d\Gamma}{d\cos\theta_i} = \frac{1}{2}(1 + \beta_i p \cos\theta_i)$$

angle of the top quark decay product with the top spin axis in the top quark rest frame

spin analyzing power for the decay product

$$\beta_i = \begin{cases} +1.0 & l^+ \quad \text{or} \quad \bar{d}\text{-quark} \\ -0.31 & \bar{\nu} \quad \text{or} \quad u-\text{quark} \\ -0.41 & b\text{-quark} \end{cases}$$

Motivation

partial decay width of top quark

average top quark polarization

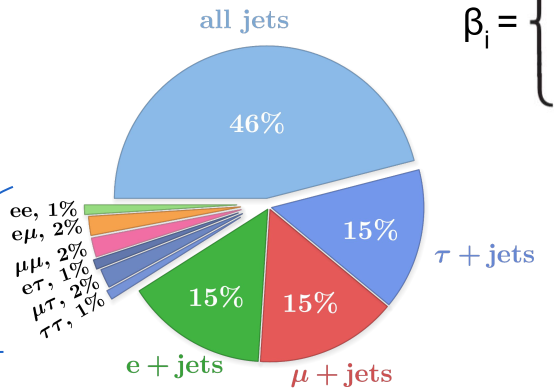
angular distribution of the decay products

$$\frac{1}{\Gamma}\frac{d\Gamma}{d\cos\theta_i} = \frac{1}{2}(1 + \beta_i p \cos\theta_i)$$

angle of the top quark decay product with the top spin axis in the top quark rest frame

spin analyzing power for the decay product

$$\beta_i = \begin{cases} +1.0 & l^+ \quad \text{or} \quad \bar{d}\text{-quark} \\ -0.31 & \bar{\nu} \quad \text{or} \quad u-\text{quark} \\ -0.41 & b\text{-quark} \end{cases}$$

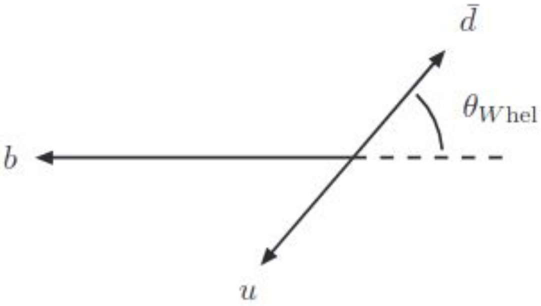


[CMS-TOP-18-006]

History of down tagging

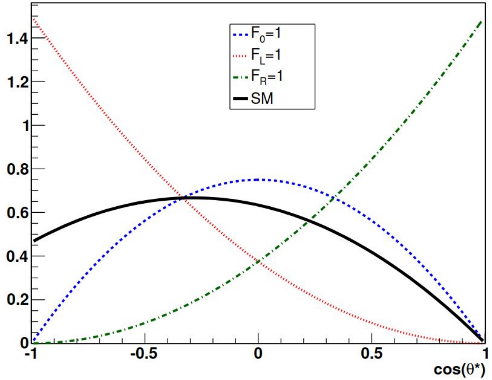
we want to maximize spin analyzing power!

"W-bosons in top decay are produced on average with negative helicity, favoring the down-type quark to be emitted closer to the b-quark"



$$f_0 \simeq \frac{m_t^2}{m_t^2 + 2m_W^2} \simeq 0.70$$

$$f_L \simeq \frac{2m_W^2}{m_t^2 + 2m_W^2} \simeq 0.30\,,$$

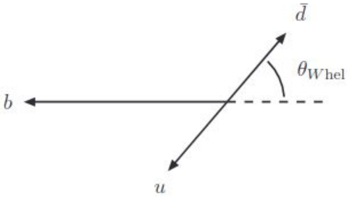


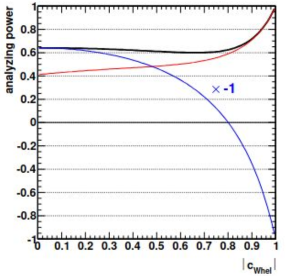
1994

History of down tagging

we want to maximize spin analyzing power!

Better Hadronic Top Quark Polarimetry





$$\vec{q}_{\mathrm{opt}}(|c_{W\mathrm{hel}}|) \equiv p(d \to q_{\mathrm{soft}})\hat{q}_{\mathrm{soft}} + p(d \to q_{\mathrm{hard}})\hat{q}_{\mathrm{hard}}$$

$$\kappa_{\mathrm{opt}}(|c_{W\mathrm{hel}}|) = |\vec{q}_{\mathrm{opt}}(|c_{W\mathrm{hel}}|)| \simeq \frac{\sqrt{s^4_{W\mathrm{hel}} m_t^2 (m_t^2 - 2m_W^2) + (1 + c^2_{W\mathrm{hel}})^2 m_W^4}}{s^2_{W\mathrm{hel}} m_t^2 + (1 + c^2_{W\mathrm{hel}}) m_W^2},$$

Spin Analyzer	Power
optimal hadronic	0.64

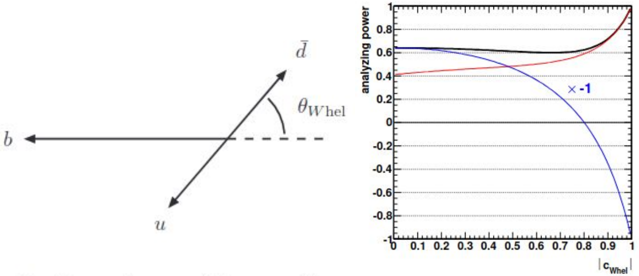
1994 **2014**

[2]Methods for measuring the charge of the progenitor quark do exist, but are not very statistically powerful for separating charge +1/3 from +2/3 (see [43]). It might nonetheless be interesting to explore what further gains could be achieved by folding in this information.

History of down tagging

we want to maximize spin analyzing power!

Better Hadronic Top Quark Polarimetry

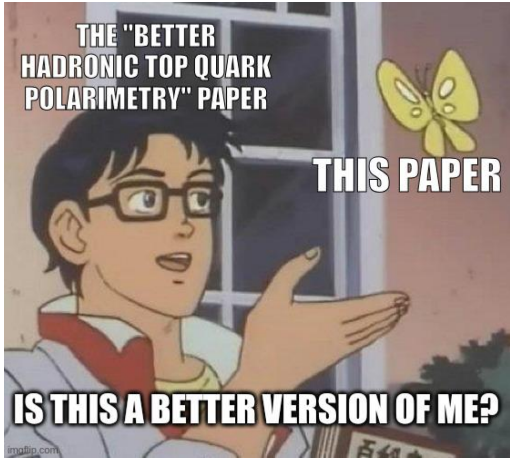


$$\vec{q}_{\rm opt}(|c_{W\rm hel}|) \equiv p(d \to q_{\rm soft})\hat{q}_{\rm soft} + p(d \to q_{\rm hard})\hat{q}_{\rm hard}$$

$$\kappa_{\rm opt}(|c_{W\rm hel}|) = |\vec{q}_{\rm opt}(|c_{W\rm hel}|)|$$

$$\simeq \frac{\sqrt{s^4_{W\rm hel} m_t^2(m_t^2 - 2m_W^2) + (1 + c^2_{W\rm hel})^2 m_W^4}}{s^2_{W\rm hel} m_t^2 + (1 + c^2_{W\rm hel}) m_W^2},$$

Spin Analyzer	Power
optimal hadronic	0.64



1994 **2014** **NOW**

[2]Methods for measuring the charge of the progenitor quark do exist, but are not very statistically powerful for separating charge +1/3 from +2/3. It might nonetheless be interesting to explore what further gains could be achieved by folding in this information.

The maths…

Fig 1.: me still trying to understand what they are doing…

including non-kinematic observables {O} other than helicity

$$\vec{q}_{\rm opt} = p(d \to q_{\rm hard}|c_W, \{\mathcal{O}\})\, \hat{q}_{\rm hard} + p(d \to q_{\rm soft}|c_W, \{\mathcal{O}\})\, \hat{q}_{\rm soft}\,.$$

…

$$\langle|\vec{q}_{\rm opt}|^2\rangle_{\mathcal{O}} = 1 - 2p(d \to q_{\rm hard}|c_W)p(d \to q_{\rm soft}|c_W)\,(1-\hat{q}_{\rm hard}\cdot\hat{q}_{\rm soft}) \times \Bigg[p(d \to q_{\rm hard}|c_W)\int d\mathcal{L}_{\mathcal{O}}\,\frac{\mathcal{L}_{\mathcal{O}}\,p(\mathcal{L}_{\mathcal{O}}|d \to q_{\rm hard})}{(p(d \to q_{\rm hard}|c_W)+\mathcal{L}_{\mathcal{O}}p(d \to q_{\rm soft}|c_W))^2} + p(d \to q_{\rm soft}|c_W)\int d\mathcal{L}_{\mathcal{O}}\,\frac{\mathcal{L}^2_{\mathcal{O}}\,p(\mathcal{L}_{\mathcal{O}}|d \to q_{\rm hard})}{(p(d \to q_{\rm hard}|c_W)+\mathcal{L}_{\mathcal{O}}p(d \to q_{\rm soft}|c_W))^2}\Bigg]\,.$$

…

$$\begin{aligned}\langle|\vec{q}_{\rm opt}|^2\rangle_{\mathcal{O}} &= 1 \qquad (15)\\ &- 2\,p(d \to q_{\rm hard}|c_W)p(d \to q_{\rm soft}|c_W)\,(1-\hat{q}_{\rm hard}\cdot\hat{q}_{\rm soft})\\ &+ 2\sigma^2\,p(d \to q_{\rm hard}|c_W)^2 p(d \to q_{\rm soft}|c_W)^2\,(1-\hat{q}_{\rm hard}\cdot\hat{q}_{\rm soft})\\ &+\cdots\,.\end{aligned}$$

The maths…

including non-kinematic observables {O} other than helicity

$$\vec{q}_{\rm opt} = p(d \to q_{\rm hard}|c_W, \{\mathcal{O}\})\, \hat{q}_{\rm hard} + p(d \to q_{\rm soft}|c_W, \{\mathcal{O}\})\, \hat{q}_{\rm soft}\,.$$

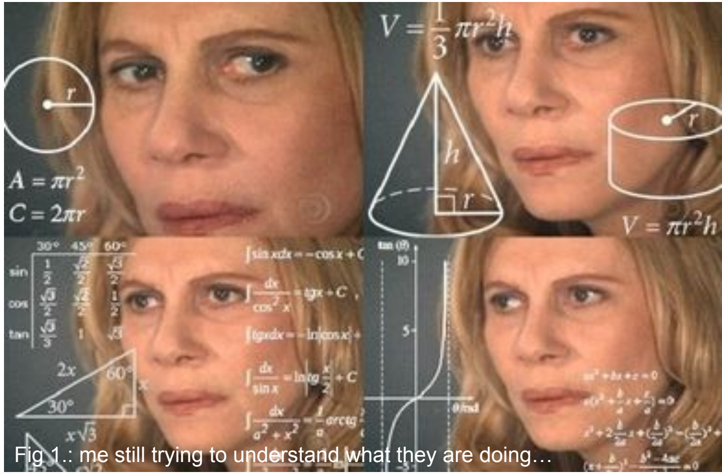


Fig 1.: me still trying to understand what they are doing…

…

$$\langle|\vec{q}_{\rm opt}|^2\rangle_{\mathcal{O}} = 1 - 2p(d \to q_{\rm hard}|c_W)p(d \to q_{\rm soft}|c_W)\,(1 - \hat{q}_{\rm hard}\cdot\hat{q}_{\rm soft}) \times \left[p(d \to q_{\rm hard}|c_W)\int d\mathcal{L}_{\mathcal{O}}\,\frac{\mathcal{L}_{\mathcal{O}}\, p(\mathcal{L}_{\mathcal{O}}|d \to q_{\rm hard})}{(p(d \to q_{\rm hard}|c_W) + \mathcal{L}_{\mathcal{O}} p(d \to q_{\rm soft}|c_W))^2} + p(d \to q_{\rm soft}|c_W)\int d\mathcal{L}_{\mathcal{O}}\,\frac{\mathcal{L}_{\mathcal{O}}^2\, p(\mathcal{L}_{\mathcal{O}}|d \to q_{\rm hard})}{(p(d \to q_{\rm hard}|c_W) + \mathcal{L}_{\mathcal{O}} p(d \to q_{\rm soft}|c_W))^2}\right].$$

…

$$\langle|\vec{q}_{\rm opt}|^2\rangle_{\mathcal{O}} = 1 - 2\,p(d \to q_{\rm hard}|c_W)p(d \to q_{\rm soft}|c_W)\,(1 - \hat{q}_{\rm hard}\cdot\hat{q}_{\rm soft}) + 2\sigma^2\, p(d \to q_{\rm hard}|c_W)^2 p(d \to q_{\rm soft}|c_W)^2\,(1 - \hat{q}_{\rm hard}\cdot\hat{q}_{\rm soft}) + \cdots\,. \tag{15}$$

Term is strictly positive

depends on the non-kinematic observables {O}

The maths…

including non-kinematic observables {O} other than helicity

$$\boxed{\sqrt{\langle |\vec{q}_{\rm opt}|^2 \rangle} \approx 0.643 + 0.163\sigma^2 + \cdots .}$$

σ^2 is the variance of $p(\mathcal{L}_{\mathcal{O}}|d \to q_{\rm hard})$

$$\mathcal{L}_{\mathcal{O}} \equiv \frac{p(\{\mathcal{O}\}|d \to q_{\rm soft})}{p(\{\mathcal{O}\}|d \to q_{\rm hard})} .$$

Analysis

Samples:

- semileptonic ttbar samples at 14 TeV

(left-, right-, and unpolarized top quarks)

- gen-level top p_T > 200 GeV

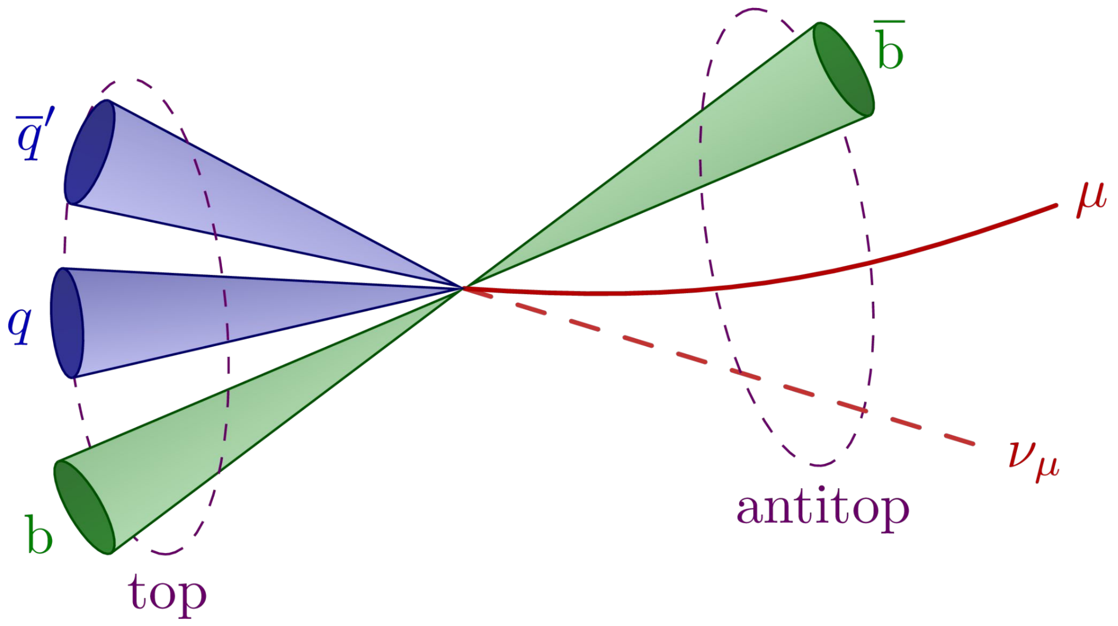
Objects:

- all particles (|η| < 3 and pT > 1 GeV)
- CA fatjet (R=1.5) with pT > 250 GeV
- declustering the fatjet

  → ≥3 subjets (m_{subjet} < 30 GeV)

Selection:

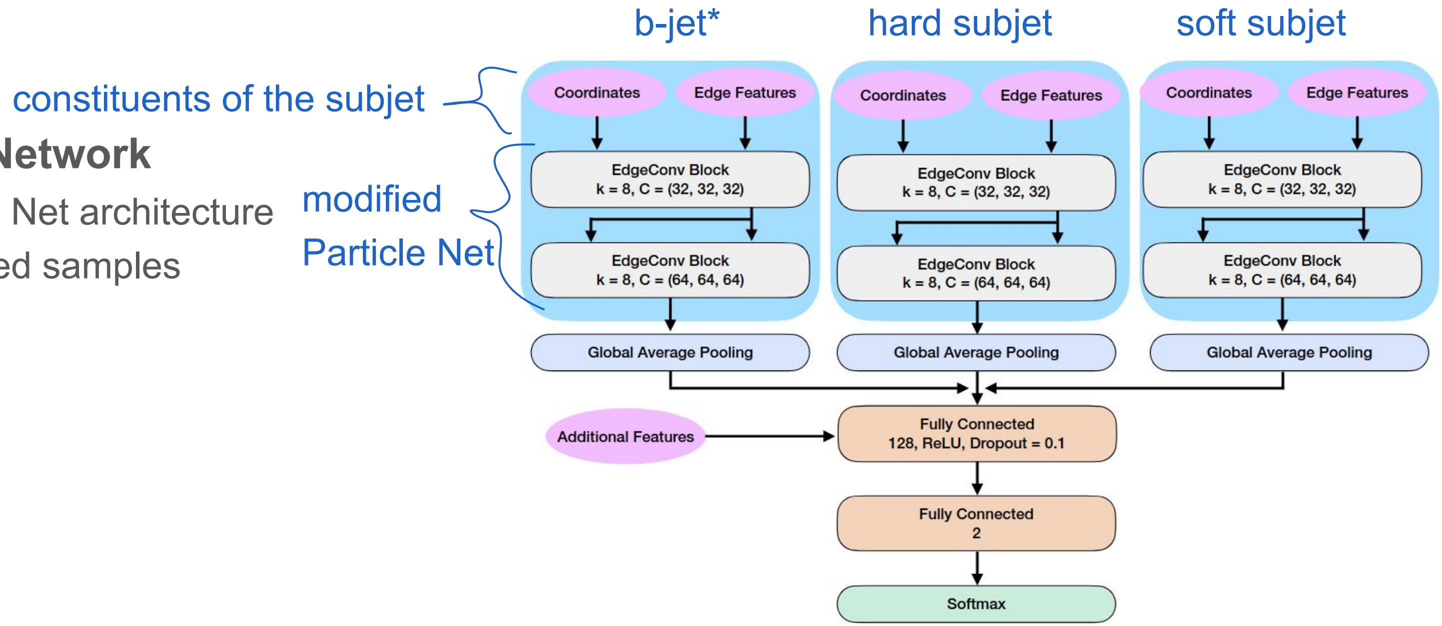
- hardest 4 subjets
- top quark reconstruction with 3rd or 4th jet, $m(j_1,j_2,j_{3/4}) \in [165, 190]$ GeV
- parton matching



Analysis

Fancy Graph Neural Network

- modified the Particle Net architecture
- trained on unpolarized samples



*In no way they state how they tag the b-jet… (I assume gen-level information and one should keep in mind the b-tagging efficiency when applying on reco level?)

Analysis

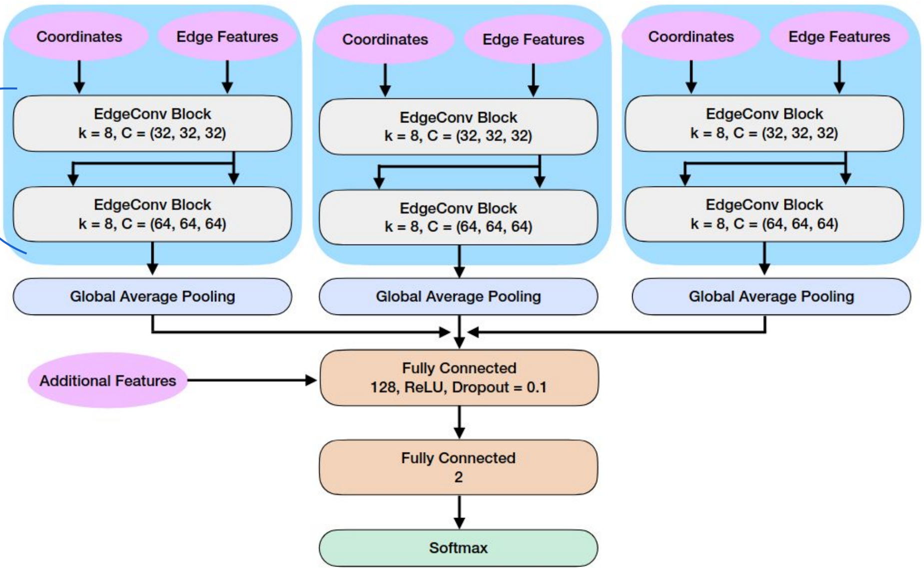
constituents of the subjet

Fancy Graph Neural Network

- modified the Particle Net architecture
- trained on unpolarized samples

modified Particle Net

b-jet* hard subjet soft subjet



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

TABLE I: Input features in GNN: the positions of points in the graph (top) and the attributes of each particle (bottom).

*In no way they state how they tag the b-jet… (I assume gen-level information and one should keep in mind the b-tagging efficiency when applying on reco level?)

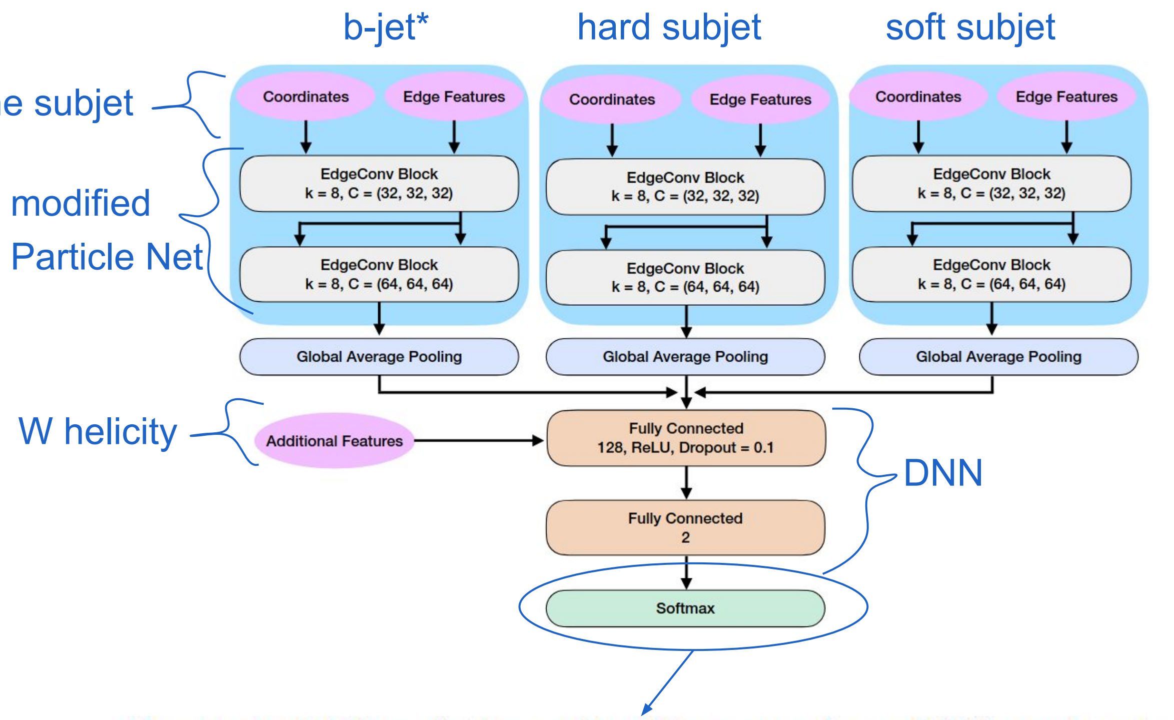
Analysis

Fancy Graph Neural Network

- modified the Particle Net architecture
- trained on unpolarized samples

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

TABLE I: Input features in GNN: the positions of points in the graph (top) and the attributes of each particle (bottom).



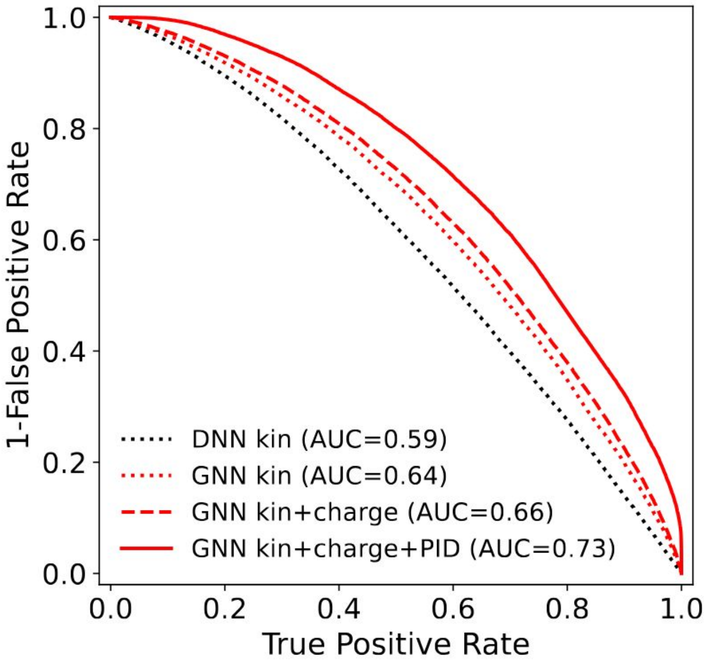
the probability of the soft $p(d \to q_{\text{soft}}|c_W, \{\mathcal{O}\})$ or hard subjet $p(d \to q_{\text{hard}}|c_W, \{\mathcal{O}\})$ being the down-type jet. In

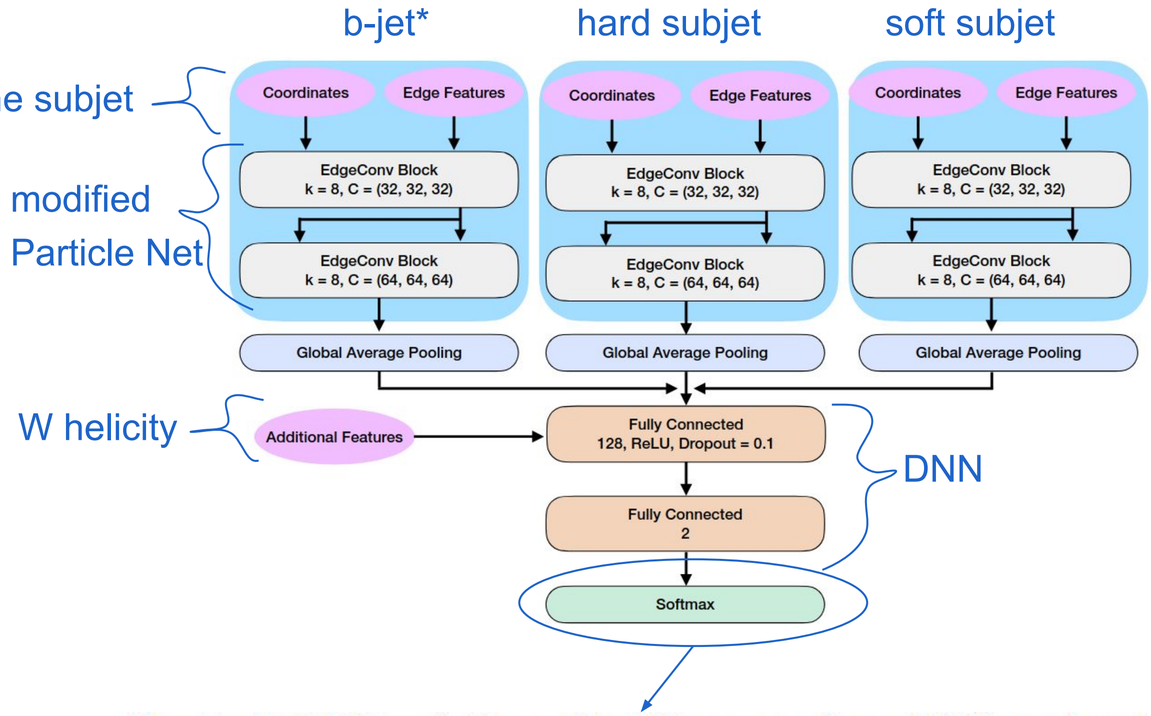
*In no way they state how they tag the b-jet… (I assume gen-level information and one should keep in mind the b-tagging efficiency when applying on reco level?)

Analysis

Fancy Graph Neural Network

- modified the Particle Net architecture
- trained on unpolarized samples





the probability of the soft $p(d \to q_{\text{soft}}|c_W, \{\mathcal{O}\})$ or hard subjet $p(d \to q_{\text{hard}}|c_W, \{\mathcal{O}\})$ being the down-type jet. In

*In no way they state how they tag the b-jet… (I assume gen-level information and one should keep in mind the b-tagging efficiency when applying on reco level?)

Analysis

b-jet* hard subjet soft subjet

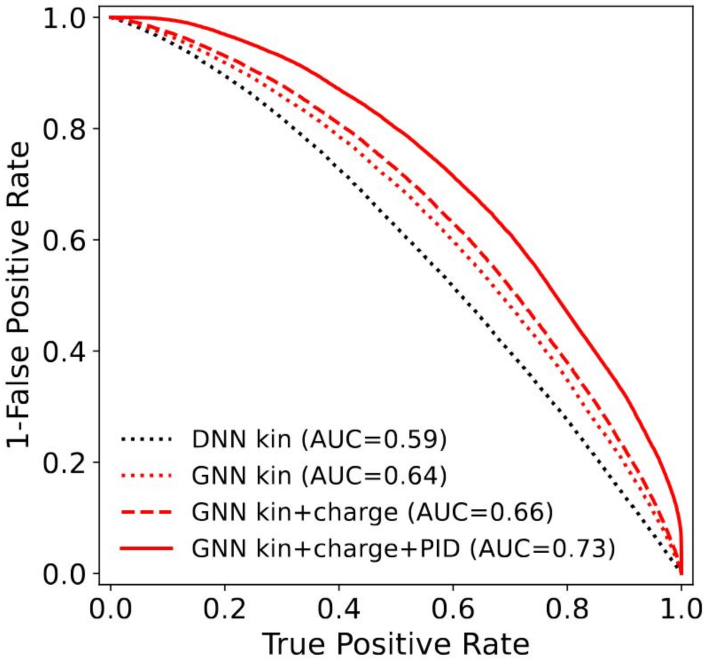
constituents of the su

Fancy Graph Neural Network

- modified the Particle Net architecture
- trained on unpolarized samples

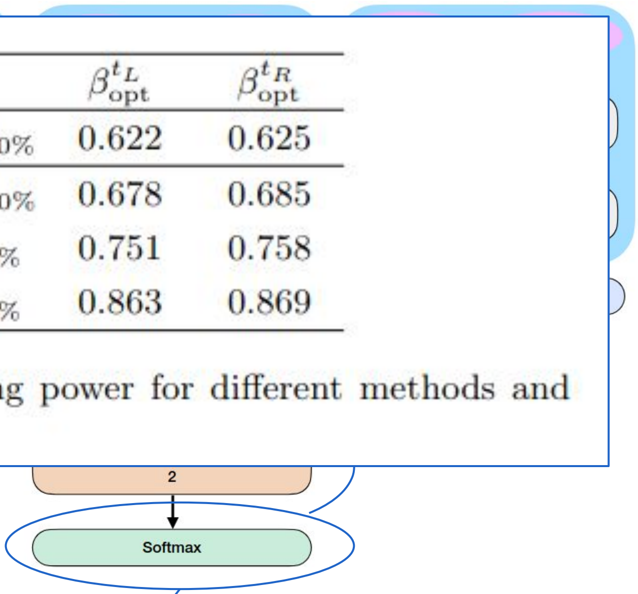
mo

Par



	$\beta^{t_L}_{\text{opt}}$	$\beta^{t_R}_{\text{opt}}$
$\text{DNN}_{\text{Eff}=100\%}$	0.622	0.625
$\text{GNN}_{\text{Eff}=100\%}$	0.678	0.685
$\text{GNN}_{\text{Eff}=50\%}$	0.751	0.758
$\text{GNN}_{\text{Eff}=20\%}$	0.863	0.869

TABLE II: Spin analyzing power for different methods and efficiencies.



the probability of the soft $p(d \to q_{\text{soft}}|c_W, \{\mathcal{O}\})$ or hard subjet $p(d \to q_{\text{hard}}|c_W, \{\mathcal{O}\})$ being the down-type jet. In

$$\vec{q}_{\text{opt}} = p(d \to q_{\text{hard}}|c_W, \{\mathcal{O}\})\, \hat{q}_{\text{hard}} + p(d \to q_{\text{soft}}|c_W, \{\mathcal{O}\})\, \hat{q}_{\text{soft}}\,.$$

Conclusion

down tagging in top decays

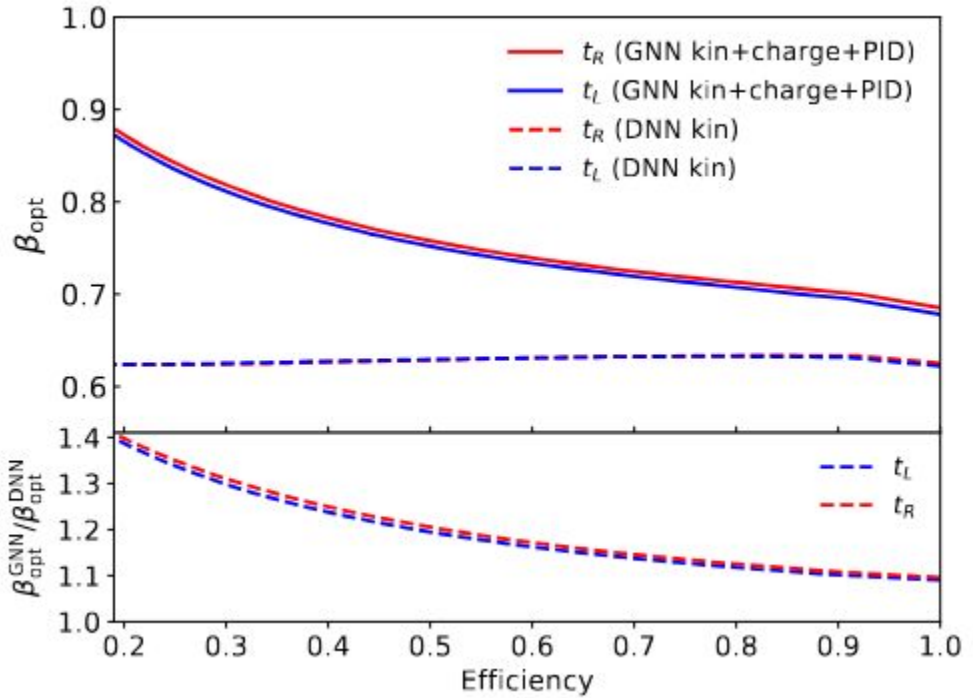
- GNN for tagging
- jet substructure kinematics and PID most important for down tagging

top quark polarization

- tagged down-type jet can be used as a proxy for top quark spin in ttbar semileptonic
- boost the top quark precision physics studies

Future Outlook?

- incorporate charm tagging since half of the hadronic top decays to charm quarks.
- actually be used in a measurement



paper link

journal club
06 Dec 2024

Hadronic Top quark polarimetry with particleNet

“playing tag with the down quark at the ttbar playground”

Jules Vandenbroeck

GHENT UNIVERSITY

CMS

Appendix

W's polarization causes $c_{W\mathrm{hel}}$ to be distributed as

$$\rho\left(c_{W\mathrm{hel}}\right) \equiv \frac{3}{8} f_R \left(1 + c_{W\mathrm{hel}}\right)^2 + \frac{3}{4} f_0 \left(1 - c_{W\mathrm{hel}}^2\right) + \frac{3}{8} f_L \left(1 - c_{W\mathrm{hel}}\right)^2 , \tag{1}$$

3

where f_R, f_0, and f_L are respectively the fractions of right-handed helicity, zero helicity, and left-handed helicity W bosons in top-frame. In the $V - A$ electroweak theory, f_R is nearly zero, and

$$\begin{aligned} f_0 &\simeq \frac{m_t^2}{m_t^2 + 2m_W^2} \simeq 0.70 \\ f_L &\simeq \frac{2m_W^2}{m_t^2 + 2m_W^2} \simeq 0.30 \,, \end{aligned} \tag{2}$$

in the approximation $m_b = 0$ and taking $m_t = 172$ GeV. By the approximate CP-invariance of the decay, anti-tops have a nearly identical distribution.

Appendix

For the DNN implementation, we input the four-momenta as (p_T, η, ϕ, E) of the b-jet, harder jet, and softer jet in that order. Additionally, the helicity angle in the top rest frame is included as an input feature, resulting in a total of 13 input features. The binary label given to each event indicates whether the harder jet corresponds to the down-type jet. The network architecture consists of three hidden layers, each with 32 dimensions with RELU activation function. The output layer is one-

Table I. After the Edge Convolution blocks, each graph is pooled and flattened in the same manner, then concatenated into a single linear input of total dimension of 192. The helicity angle is included as a supplementary feature, by concatenating it with the linear layer following the edge convolutions. The combined inputs are then fed into a fully connected linear layer with 128 neurons before the output layer. The network architecture is depicted in Fig. 1. The subjets' input order and labeling are consistent with the DNN case. Both networks are optimized using the Adam optimizer.

W's polarization causes $c_{W\text{hel}}$ to be distributed as

$$\rho\left(c_{W\text{hel}}\right) \equiv \frac{3}{8} f_R \left(1 + c_{W\text{hel}}\right)^2 + \frac{3}{4} f_0 \left(1 - c_{W\text{hel}}^2\right) + \frac{3}{8} f_L \left(1 - c_{W\text{hel}}\right)^2 , \qquad (1)$$

3

where f_R, f_0, and f_L are respectively the fractions of right-handed helicity, zero helicity, and left-handed helicity W bosons in top-frame. In the $V - A$ electroweak theory, f_R is nearly zero, and

$$\begin{aligned} f_0 &\simeq \frac{m_t^2}{m_t^2 + 2m_W^2} \simeq 0.70 \\ f_L &\simeq \frac{2m_W^2}{m_t^2 + 2m_W^2} \simeq 0.30 \, , \end{aligned} \qquad (2)$$

in the approximation $m_b = 0$ and taking $m_t = 172$ GeV. By the approximate CP-invariance of the decay, anti-tops have a nearly identical distribution.