Spectral Analysis of Jet Substructure with Neural Network: Boosted Higgs Case

Sung Hak Lim

KEK

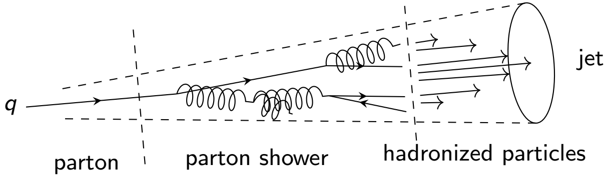
Supplementary materials for BOOST 2018, Paris, France

Jul. 2018

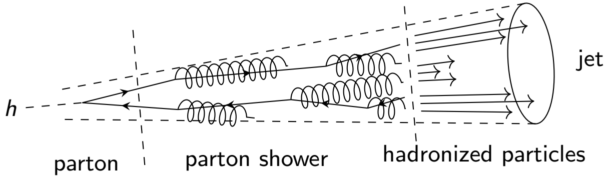
based on **S. H. Lim**, M. M. Nojiri, arXiv:1807.03312

Jets and Boosted Particles

- A jet is a collimated cluster of particles. It is often produced from a colored parton.



- As LHC stacking up multi *TeV* center-of-mass energy events, boosted heavy particles can be produced and form a single collimated cluster of particles similar to the QCD jet. ($m_{EW}/\sqrt{\hat{s}} = \mathcal{O}(0.1)$)



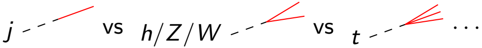
- We have to differentiate these non-QCD jets from QCD jets to maximize sensitivity of channels involving boosted particles.
- We use substructure observables on this purpose.

Conventional observables of jet substructures

There are lots of observables of jet substructures, which are focussing on a particular substructure. A few examples are:

- mass of the mother particle: m_{jet}, trimmed m_{jet}, $\cdots$

- n-prong jet: n-subjettiness ratio, D_2, $\cdots$



- subjet p_T asymmetry: mass drop tagger, $\cdots$

- color charge: jet girth (a kind of width), $\cdots$

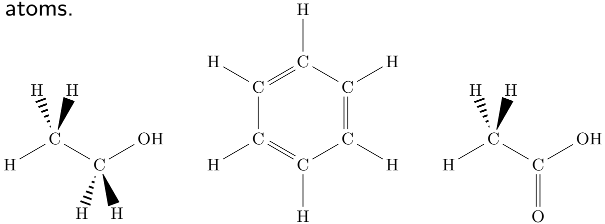
- color substructures: jet pull between subjets, $\cdots$

Q: is there any generic framework that unifies these variables?

- I want to introduce an analogy from Organic Chemistry.

Organic Molecules

- Organic molecules are complex molecules contains carbon, hydrogen, and other atoms.



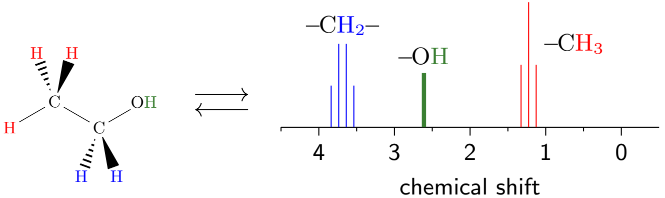
Ethanol Benzene Acetic Acid

- How chemists identified these complex substructures?

  Proton Nuclear Magnetic Resonance (^{1}H-NMR) Spectroscopy

Analogy of Jet Substructures and Organic Molecules

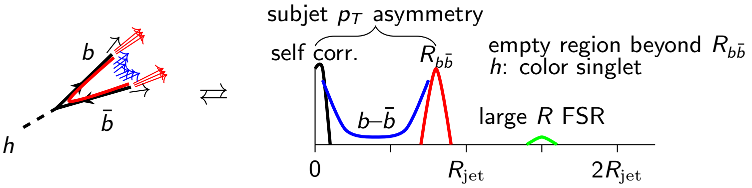
- ^{1}H-NMR is a successful example of spectral analysis of molecular substructures.



- The perturbative description of jets is analogous to organic molecules.

Spectral analysis	^{1}H-NMR spectroscopy	$\rightarrow$	??
Primary geometry	Carbon skeleton	$\rightarrow$	partons from the decay
Observables	Hydrogen	$\rightarrow$	hadronized particles

- Can we build a similar analysis framework of jet substructures?



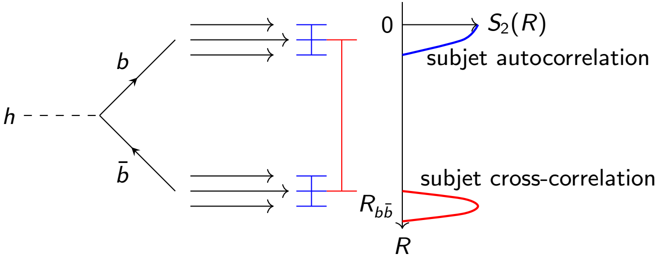
- We will introduce a spectral analysis of jet substructure.

A spectral function of jet substructure

- We define a binned spectral function of transverse momenta of all particle pairs i and j, $p_{T,i}$ and $p_{T,j}$, and their angular distance $R_{ij} = \sqrt{\eta_{ij}^2 + \phi_{ij}^2}$,

$$S_2(R;\Delta R) = \frac{1}{\Delta R} \sum_{\substack{i,j\in \text{jet} \\ R_{ij}\in[R,R+\Delta R)}} p_{T,i}\, p_{T,j}, \qquad (1)$$

- This spectral function observes two-point correlations in a jet.



$$S_2(R) = \left(p_{T,b}^2 + p_{T,\bar{b}}^2\right)\cdot\delta(R) + 2p_{T,b}\,p_{T,\bar{b}}\cdot\delta(R - R_{b\bar{b}}).$$

- Hence, this spectral function contains non-local correlation in jets. The correlations can be used for detailed jet substructure studies.

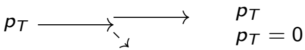
Physical Interpretation of $S_2(R)$

$$\int_0^\infty dR\, S_2(R) = \left(\sum_{i\in\text{jet}} p_{T,i}\right)^2 \approx p_{T,\text{jet}}^2, \quad (2)$$

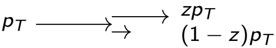
$$\int_0^\infty dR\, R^2 S_2(R) = \sum_{i,j\in\text{jet}} p_{T,i} p_{T,j} R_{ij}^2 \approx 2m_{\text{jet}}^2. \quad (3)$$

IRC safety of $S_2(R)$

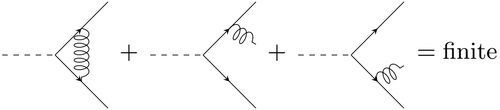
- The $S_2(R)$ spectrum is better to be infrared and collinear (IRC) safe, namely invariant under soft and collinear radiations.
  - Soft radiation: the parton radiates a $p_T = 0$ parton.



  - Collinear radiation: the parton splits in same direction $\vec{R}$.



- Otherwise, the virtual and real corrections in higher order perturbation theory do not sum up, and such a IRC unsafe observables are hard to be estimated from perturbative QCD calculations. (KNL theorem)



(4)

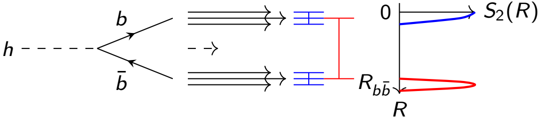
- $S_2(R)$ is IRC safe because soft or collinear radiation does not introduce any new angular scale R.

IRC safety of $S_2(R)$

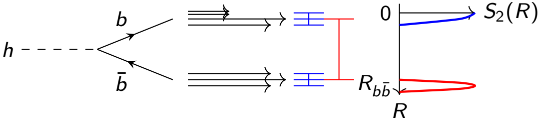
- The $S_2(R)$ spectrum is IRC safe.

$$S_2(R;\Delta R) = \frac{1}{\Delta R} \sum_{\substack{i,j\in \text{jet} \\ R_{ij}\in[R,R+\Delta R)}} p_{T,i} p_{T,j}, \quad (5)$$

  - Soft radiation is ignored safely because $p_{T,i}p_{T,j} = 0$ if i or j is a soft radiation.

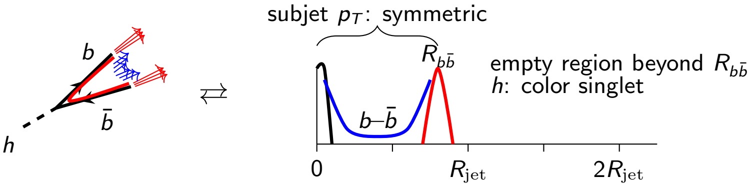


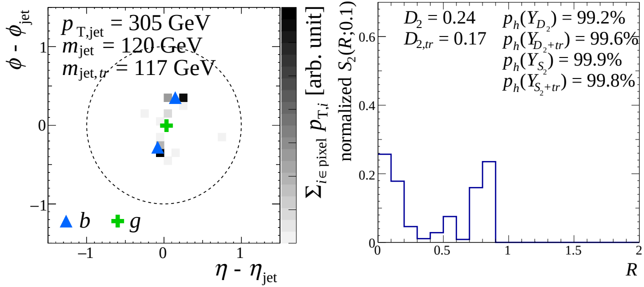
  - Collinear radiation is okay because $\vec{R}$ does not change and the radiation products always sum up.



- Now, let's check how $S_2(R)$ behaves with more realistic events at the detector level.

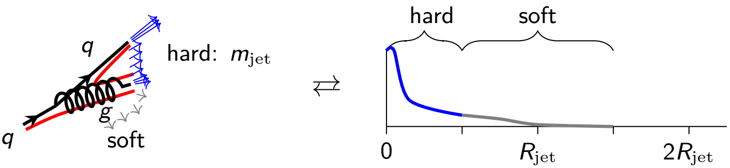
A typical Higgs jet

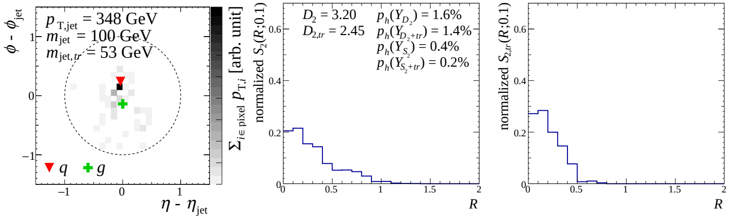




$$S_2(R) = \left(p_{T,b}^2 + p_{T,\bar{b}}^2\right) \cdot \delta(R) + 2p_{T,b}p_{T,\bar{b}} \cdot \delta(R - R_{b\bar{b}}).$$

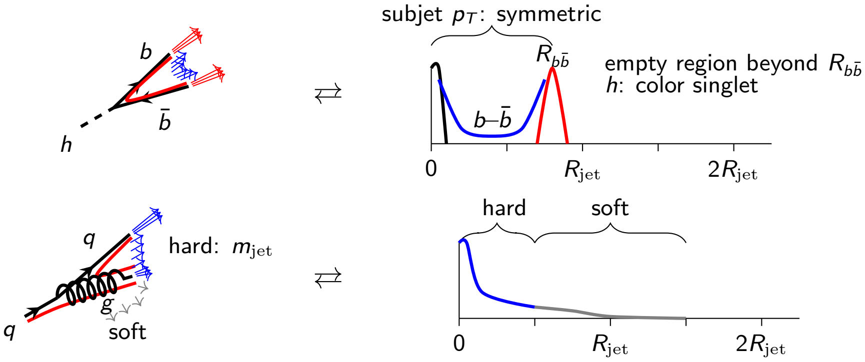
A typical QCD jet and jet trimming





- Jet trimming removes soft subjets in the jet.
- Jet trimming helps differentiating hard and soft jet substructures.

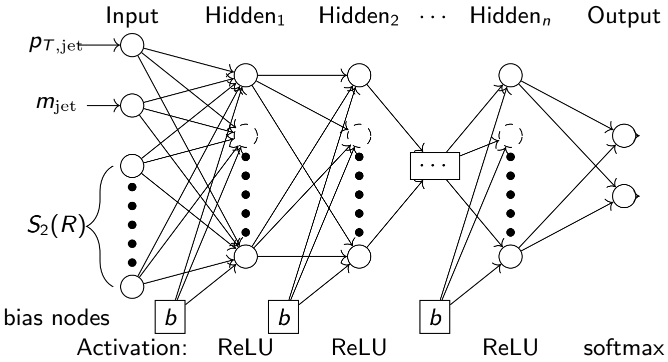
Differntiating Higgs jets and QCD jets



- $S_2(R)$ contains various information on jet substructures. But how can we quantify it for classification?

Deep Learning

Artificial Neural Network



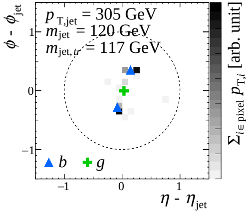
- The artificial neural network is a mathematical model of functions motivated from a biological neural network.
- To make a long story short, the neural network is a function maps inputs to outputs having lots of internal parameters needed to be optimized.

$$ANN(\{x_i\}) = f^{(n)}(W^{(n)} \cdots f^{(2)}(W^{(2)}_{jk} f^{(1)}(W^{(1)}_{ij} x_i + b^{(1)}_j) + b^{(2)}_k) \cdots + b^{(n)})$$

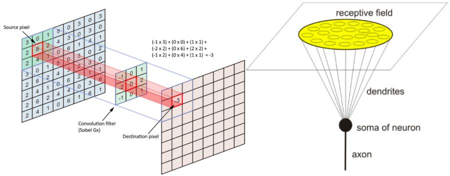
- The network setup for jet substructure analysis depends on how you interpret jets.

Jet as an Image

- We may interpret the calorimeter energy deposit as an image, and apply image recongition techniques.



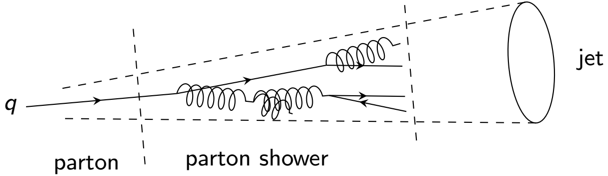
- Convolutional neural network



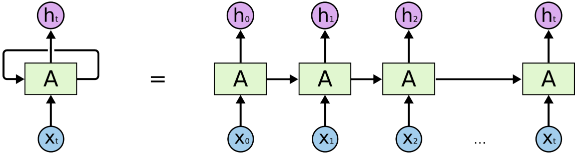
  - This ANN mimicks human eye and can be used for image recongnition.
  - It focuses on local spatial correlations
- References: 1407.5675, 1511.05190

Jet as a sequential data

- The parton shower description of jets can be nterpreted as a sequential data.

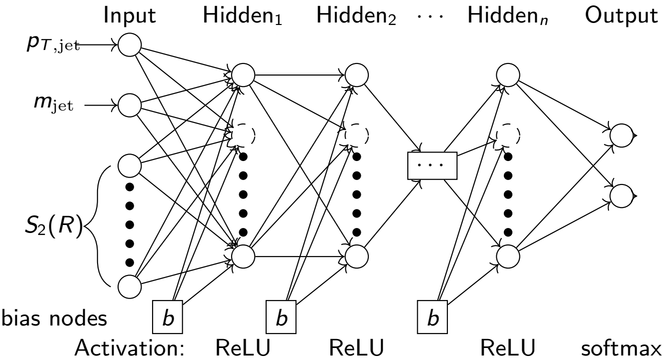


- We may recluster the jet to get a parton shower history and feed this to ANN.
- Recurrent Nerual Network



  - The ANN understand data in sequence and find out correlations by reading data sequentially.
  - Focussing on local temporal correlations
- References: 1607.08633, 1702.00748

Our ANN setup



- We used a simple shallow fully-connected neural network.
- Inputs for $S_2(R)$ analysis
  - $\mathcal{N}_{S_2}$: $\{x_i\}_{S_2} = \{p_{T,\mathrm{jet}}, m_{\mathrm{jet}}, S_2(0;0.1), \cdots, S_2(1.9;0.1)\}$
  - $\mathcal{N}_{S_2+tr}$: $\{x_i\}_{S_2+tr}$
    $= \{x_i\}_{S_2} \cup \{p_{T,\mathrm{jet},tr}, m_{\mathrm{jet},tr}, S_{2,tr}(0;0.1), \cdots, S_{2,tr}(1.9;0.1)\}$

Comparison to another variable

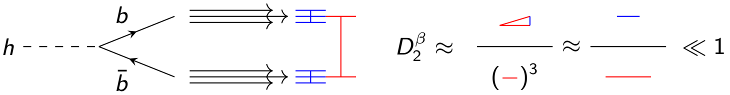
- For comparison, we prepare an ANN anlysis with D_2 variable (1409.6298)

$$e_2^\beta = \frac{1}{p_{T,\mathrm{jet}}^2} \sum_{\substack{i,j\in \mathrm{jet} \\ i<j}} p_{T,i} p_{T,j} R_{ij}^\beta, \tag{6}$$

$$e_3^\beta = \frac{1}{p_{T,\mathrm{jet}}^3} \sum_{\substack{i,j,k\in \mathrm{jet} \\ i<j<k}} p_{T,i} p_{T,j} p_{T,k} R_{ij}^\beta R_{jk}^\beta R_{ki}^\beta, \tag{7}$$

$$D_2^\beta = \frac{e_3^\beta}{(e_2^\beta)^3} \sim \frac{\triangle}{(-)^3}, \tag{8}$$

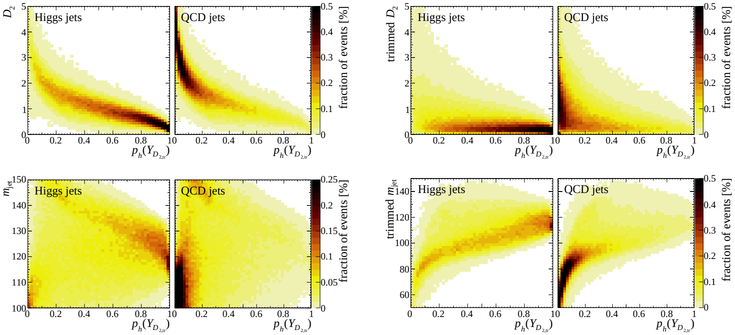
- For Higgs jets, D_2 is small because three-point energy correlation have to count soft or collinear radiation.



- Inputs for ANN with D_2
  - $\mathcal{N}_{D_2}$: $\{x_i\}_{D_2} = \{p_{T,\mathrm{jet}}, m_{\mathrm{jet}}, D_2^{\beta=2}\}$
  - $\mathcal{N}_{D_2+tr}$: $\{x_i\}_{D_2+tr} = \{x_i\}_{D_2} \cup \{p_{T,\mathrm{jet},tr}, m_{\mathrm{jet},tr}, D_{2,tr}^{\beta=2}\}$

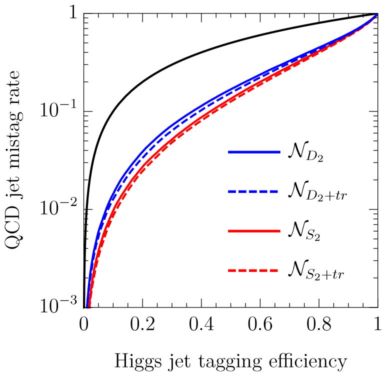
Performance check of ANN with D_2

- We first check that the performance of $\mathcal{N}_{D_2+tr}$ with The distributions of Higgs jets and QCD jets in the Higgs-like probability $p_h(Y_{D_2+tr})$ and inputs in $\{x_i\}_{D_2+tr}$.



- $\mathcal{N}_{D_2+tr}$ has a tendency to classify jets with small D_2 and $m_{\rm jet} \approx 125$ GeV as a Higgs jet. This behavior is similar to the conventional cut-based analysis.
- Now it looks working well, so let us compare it with $\mathcal{N}_{S_2+tr}$.

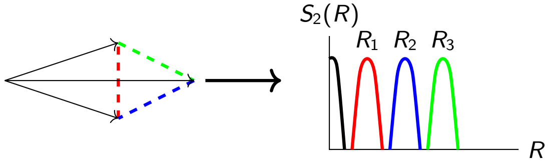
ROC curve



- Event preselection:
  - $m_{\rm jet} \in [100, 150]$ GeV.
  - $p_{T,{\rm jet}} \in [300, 400]$ GeV.
  - Then $R_{b\bar{b}} \gtrsim 0.6$ and anti-k_T algorithm with $R_{\rm J} = 1$ finds the collimated cluster well.
  - For Higgs jets, at least one b parton should be found in the jet.
- At the Higgs tagging efficiency 0.4 (0.2), QCD jet mistag rate of $\mathcal{N}_{S_2+tr}$ is reduced by 21.6% (26.7%) compared to that of $\mathcal{N}_{D_2+tr}$.
- Why $S_2(R)$ is doing better than D_2?

$S_2(R)$ and D_2

- $S_2(R)$ includes D_2 partially. Let's consider a three-prong jet.



  - Two-point correlation function.

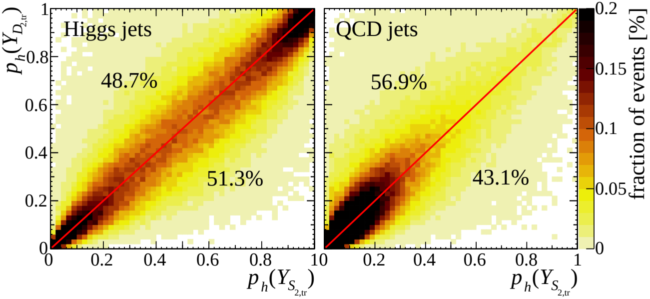
$$\int_0^\infty dR\, R^\beta S_2(R) = 2p_{T,\mathrm{jet}}^2 e_2^\beta \tag{9}$$

  - Three-point correlation function.

$$e_3^\beta \approx p_{T,\mathrm{jet}}^{-1} \cdot \sqrt{(\Delta R)^3 S_2(R_1;\Delta R) S_2(R_2;\Delta R) S_2(R_3;\Delta R)} R_1^\beta R_2^\beta R_3^\beta \tag{10}$$

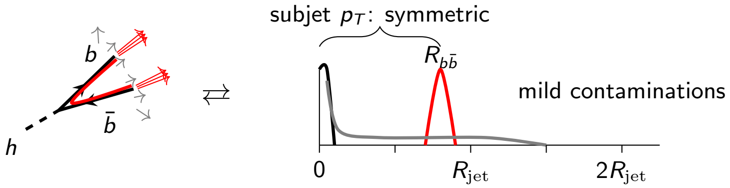
- Therefore, we may build D_2 from $S_2(R)$ if it is required.
- How the classifier $\mathcal{N}_{S_2+tr}$ is correlated to $\mathcal{N}_{D_2+tr}$ and why $\mathcal{N}_{S_2+tr}$ is doing better job?

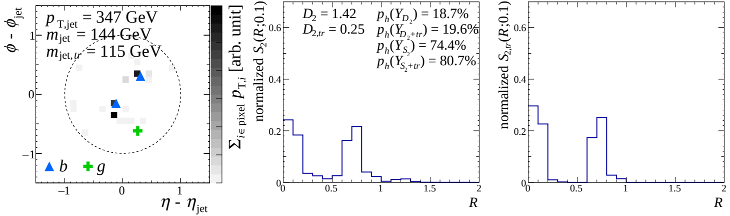
Correlation between taggers



- For Higgs jets, the lower triangular region contains more events compared with the upper triangular region, 51.3% of the total events.
- For QCD jets, the lower triangular region contains less events, 43.1%.
- Hence, $\mathcal{N}_{S_2+tr}$ improves signal and background ratio S/B from $\mathcal{N}_{D_2+tr}$.
- Let's check what kinds of jets are located in the off-diagonal region.

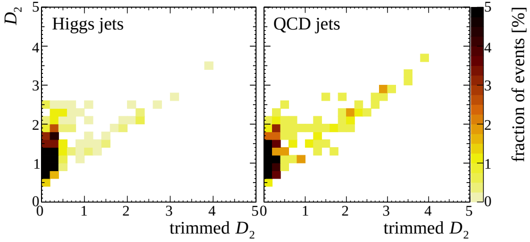
A jet tagged in S_2 analysis only





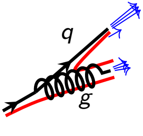
- This jet looks like a Higgs jet but why $\mathcal{N}_{D_2+tr}$ classify it as a QCD jet?

A jet tagged in S_2 analysis only

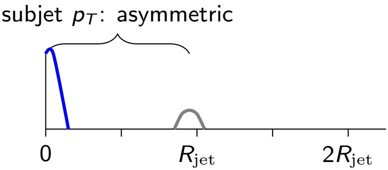


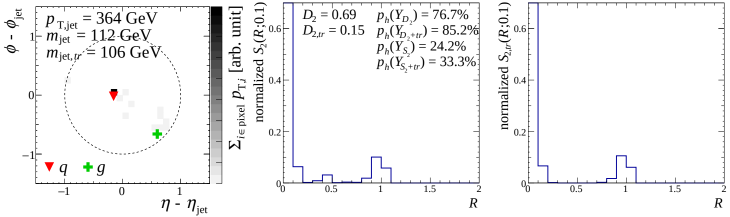
- Event selection: $p_h(Y_{D_{2,tr}}) < 30\%$ and $p_h(Y_{S_{2,tr}}) > 70\%$.
- ANN tries to utilize every information in the inputs.
- As a result, large D_2 and small trimmed D_2 are signs of large angle soft activity which is a QCD jet's feature compared to a Higgs jet. Hence $\mathcal{N}_{D_2+tr}$ classifies these jets as a QCD jets.

A jet tagged in D_2 analysis only



⇄

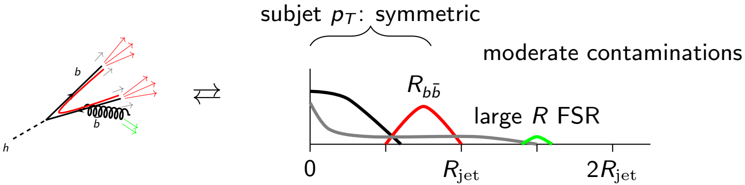


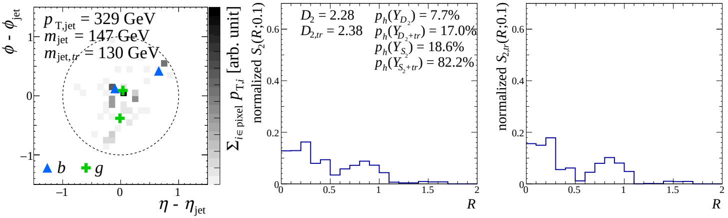


$$S_2(R) = \left(p_{T,b}^2 + p_{T,\bar{b}}^2\right) \cdot \delta(R) + 2p_{T,b}p_{T,\bar{b}} \cdot \delta(R - R_{b\bar{b}}).$$

- p_T asymmetric subjets are often occur in QCD jets.
- Even though this jet is two-prong, it's better to avoid this jet to enhance S/B ratio as mass drop tagger does.

A jet tagged in S_2 analysis with trimming only





- This jet has two-prong substructure, but the subjets are wide and contaminated by other QCD activities.
- Jet trimming helps differentiating hard and soft substructure, and $\mathcal{N}_{S_2+tr}$ classify this jet as a Higgs jet.

Conclusion

- $\mathcal{N}_{S_2}$ and $\mathcal{N}_{S_2+tr}$ learned non-local correlations in jets from the spectrum.
- $\mathcal{N}_{S_2}$ discriminates between boosted Higgs jets and QCD jets with better performance compared to $\mathcal{N}_{D_2}$.
- Introducing trimming to $S_2(R)$ helps separating hard and soft substructures, and the ANN with trimmed observable outperforms the ANN without trimming.
- The $S_2(R)$ has information on multi-point correlations and is easily applicable to other jet substructures.