Spectral Analysis of Jet Substructure with Neural Networks: Boosted Higgs and Sgluon Case

Sung Hak Lim

KEK

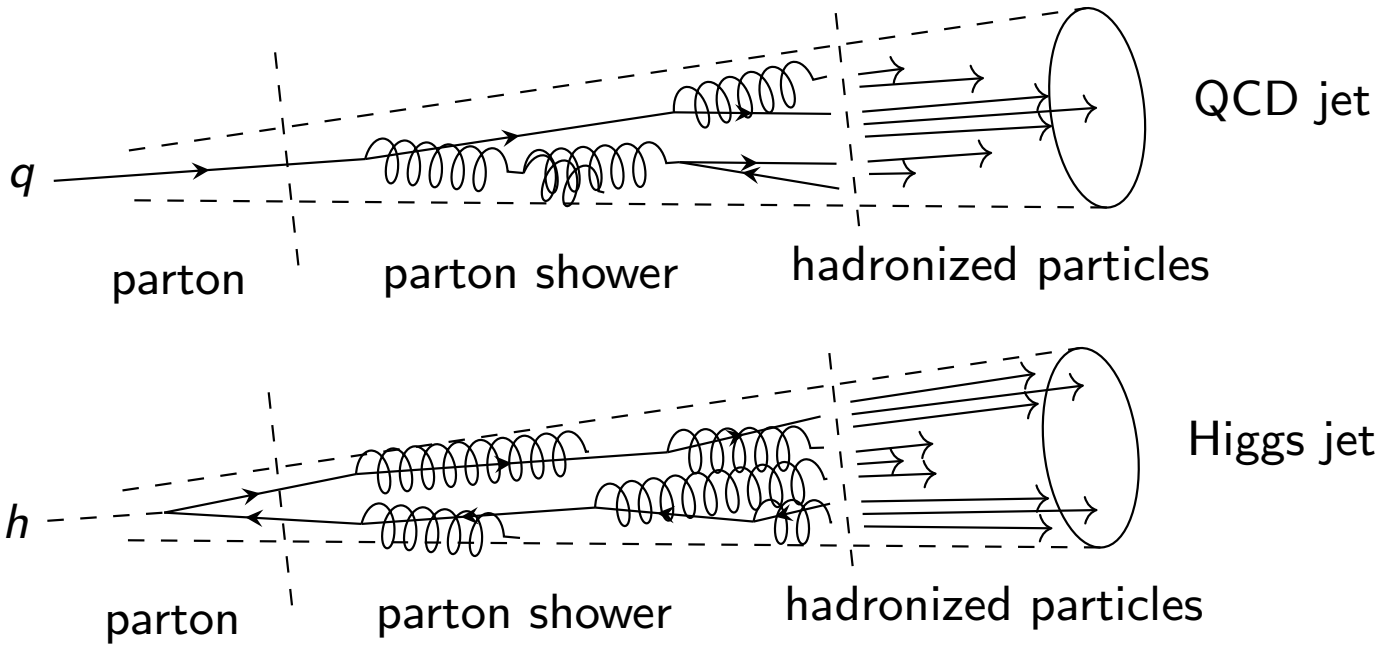
Beyond the BSM, Ikaho Onsen, Gunma, Japan

Oct. 2018

based on **S. H. Lim**, M. M. Nojiri, arXiv:1807.03312, submitted to JHEP
A. Chakraborty, **S. H. Lim**, M. M. Nojiri, work in progress

Jets and Boosted Particles

- At a hadron-hadron collider, such as LHC, a collimated particle cluster called a jet often appears when a colored parton is produced.
- As LHC stacking up multi TeV center-of-mass energy events, boosted heavy particles is produced and form a single collimated cluster of particles similar to the QCD jets. ($m_{EW}/\sqrt{\hat{s}} = \mathcal{O}(0.1)$)



- We have to differentiate these non-QCD jets from QCD jets to maximize sensitivity of channels involving boosted particles.

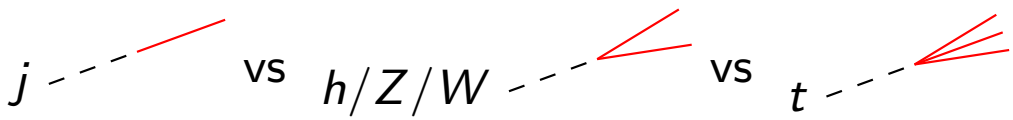
Jet Substructure !

A Catalog of Observables of Jet Substructure

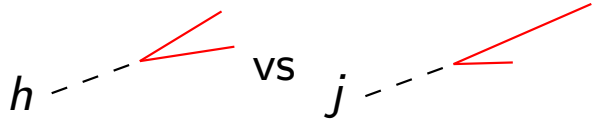
There are lots of observables of jet substructure, which are focussing on particular substructure. A few examples are:

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, $\cdots$

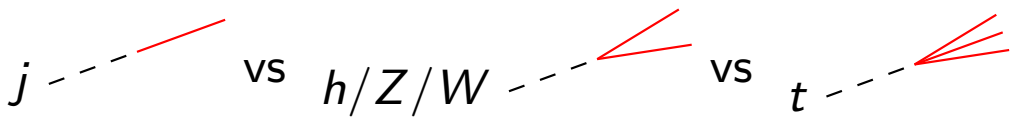
color flow:
jet pull between subjets, $\cdots$

A Catalog of Observables of Jet Substructure

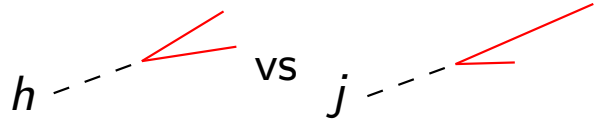
There are lots of observables of jet substructure, which are focussing on particular substructure. A few examples are:

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, $\cdots$

color flow:
jet pull between subjets, $\cdots$

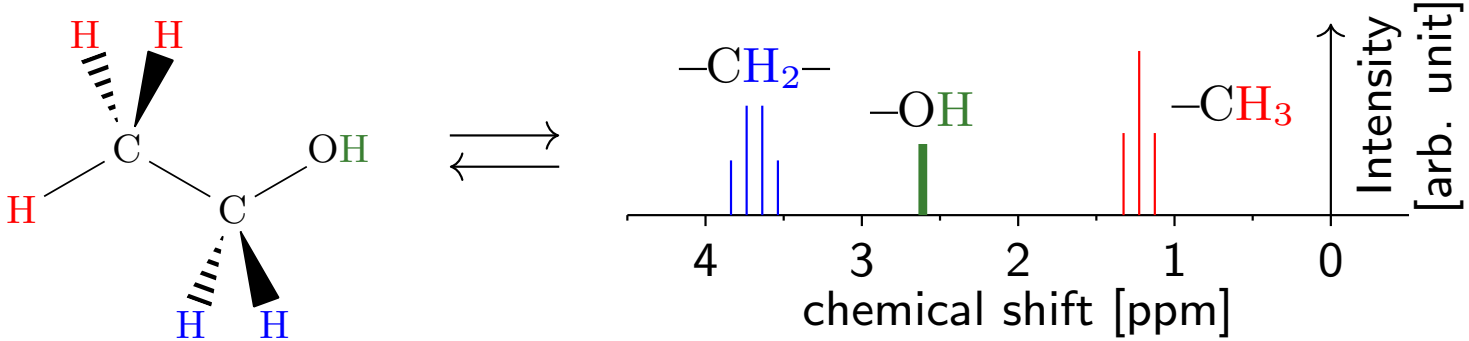
Q: is there any generic framework which is highly visual and unifies these analyses?

- I want to introduce an analogy from

  ^{1}H-NMR Analysis in Organic Chemistry.

A Spectral Function from Analogy of Jet Substructures and ^{1}H-NMR Analysis in Organic Chemistry

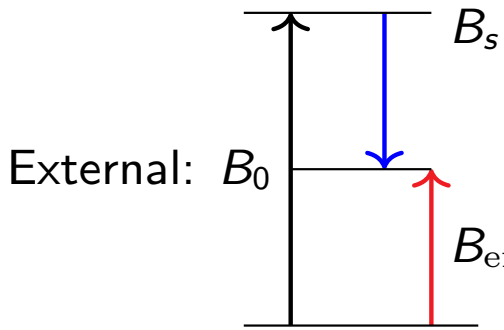
- The ^{1}H-NMR spectrum can be used for identifying the topological skeleton of an organic molecule, the carbon skeleton, from interaction between hydrogen nuclei and the rest of molecular structure.



- Chemical shift (δ): a resonant frequency of precession reletaive to a reference in a magnetic field.

$$\nu = \frac{\gamma}{2\pi} B_{\text{eff}}, \quad B_{\text{eff}} = B_0 + B_s$$

$$\delta = \frac{\nu_{\text{sample}} - \nu_{\text{ref}}}{\nu_{\text{ref}}}$$

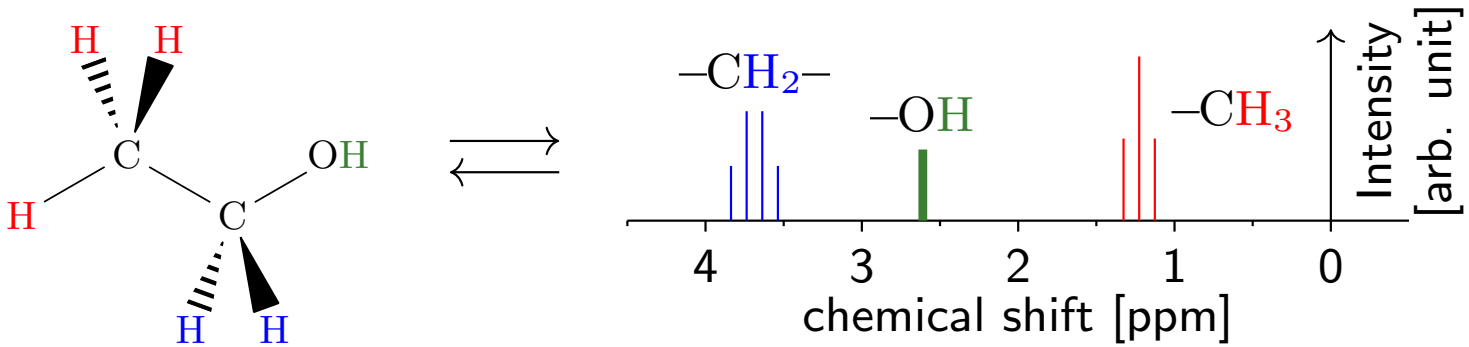


B_s : from molecular structure
-electron cloud
-nearby hydrogen spin

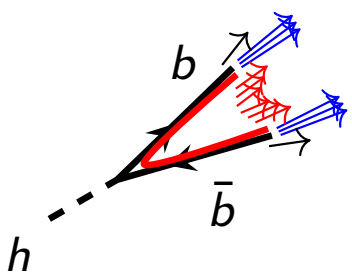
B_{eff}: nucleus sees

A Spectral Function from Analogy of Jet Substructures and ^{1}H-NMR Analysis in Organic Chemistry

- The ^{1}H-NMR spectrum can be used for identifying the topological skeleton of an organic molecule, the carbon skeleton, from interaction between hydrogen nuclei and the rest of molecular structure.

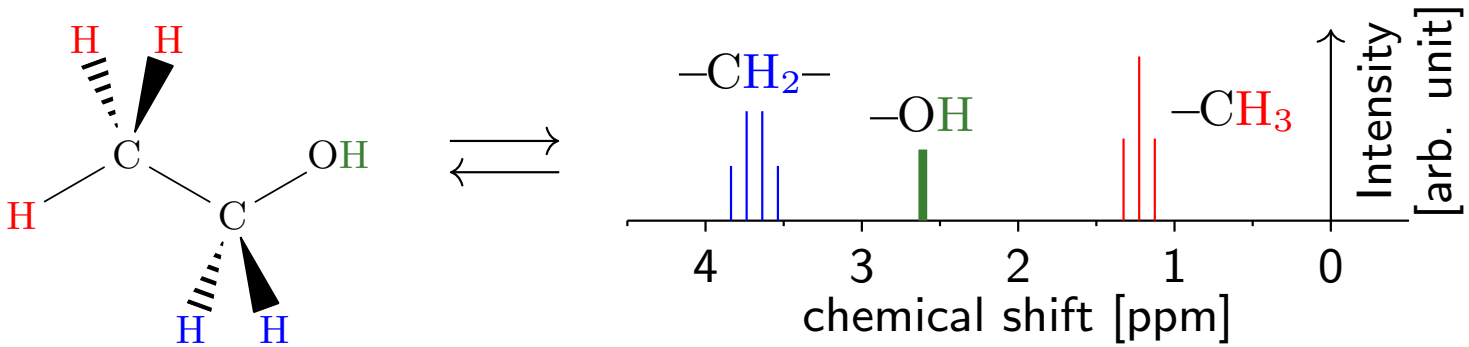


- Jets also have a topological skeleton, the originating partons.

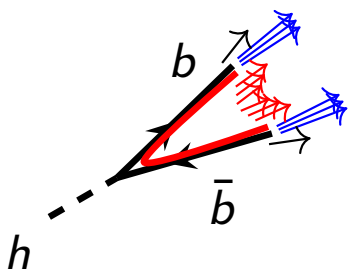


A Spectral Function from Analogy of Jet Substructures and ^{1}H-NMR Analysis in Organic Chemistry

- The ^{1}H-NMR spectrum can be used for identifying the topological skeleton of an organic molecule, the carbon skeleton, from interaction between hydrogen nuclei and the rest of molecular structure.



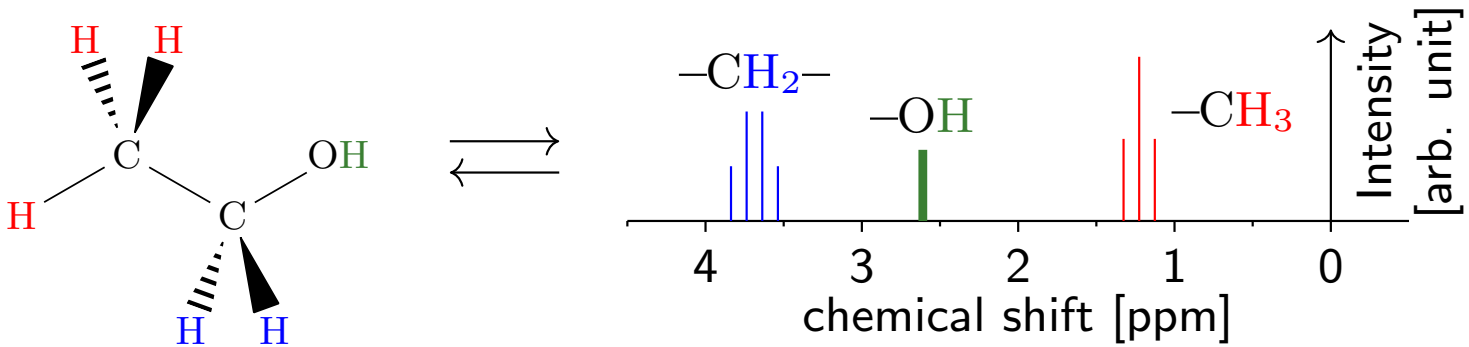
- Jets also have a topological skeleton, the originating partons.



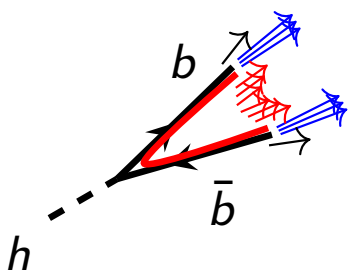
- Can we build a similar analysis framework of jet substructure to identify the originating partons of the jet?

A Spectral Function from Analogy of Jet Substructures and ^{1}H-NMR Analysis in Organic Chemistry

- The ^{1}H-NMR spectrum can be used for identifying the topological skeleton of an organic molecule, the carbon skeleton, from interaction between hydrogen nuclei and the rest of molecular structure.



- Jets also have a topological skeleton, the originating partons.

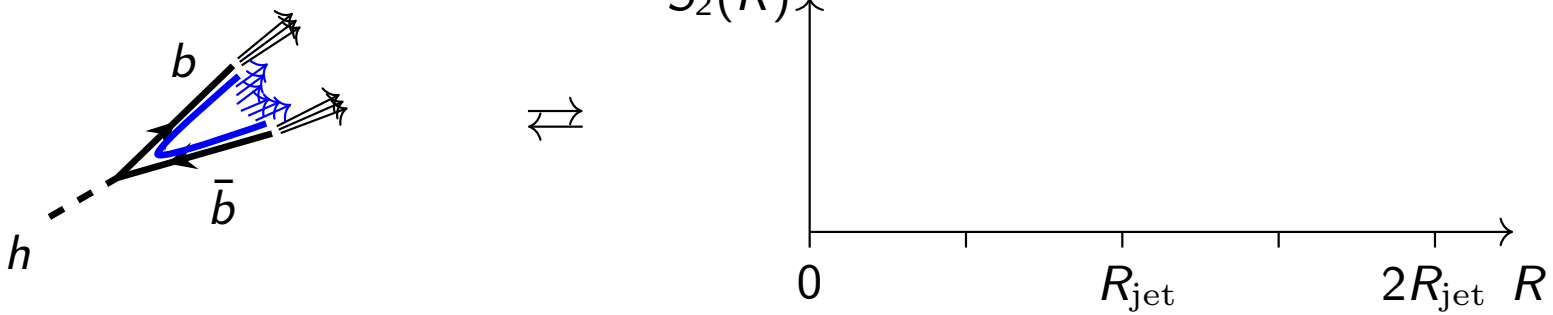


- To measure correlation between jet constituents, we define an IRC safe binned spectral function of angular distance R,

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

A typical Higgs jet

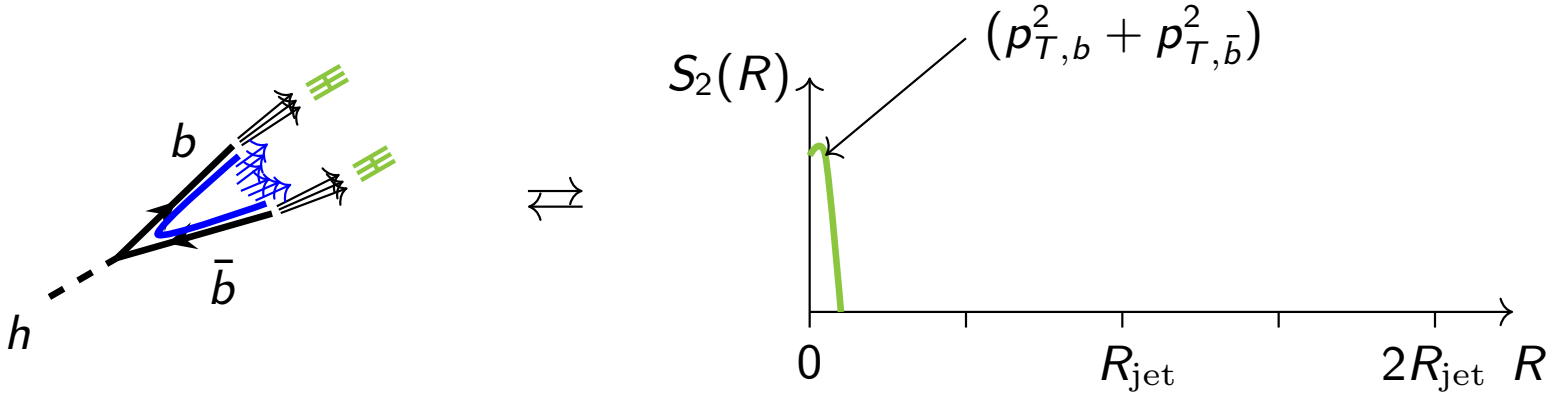
- The spectrum contains information on non-local correlation in the jet.



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

A typical Higgs jet

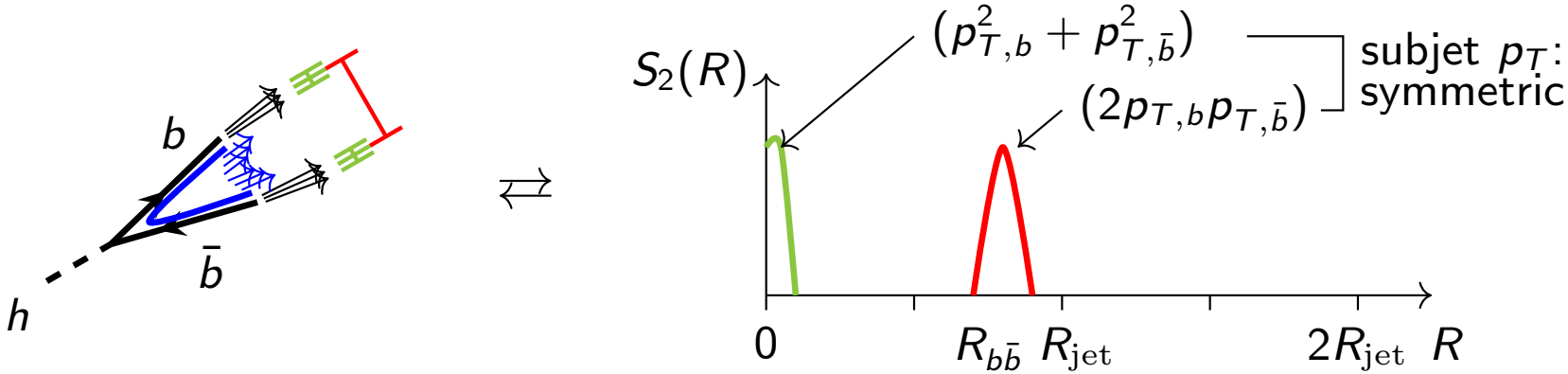
- The spectrum contains information on non-local correlation in the jet.



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

A typical Higgs jet

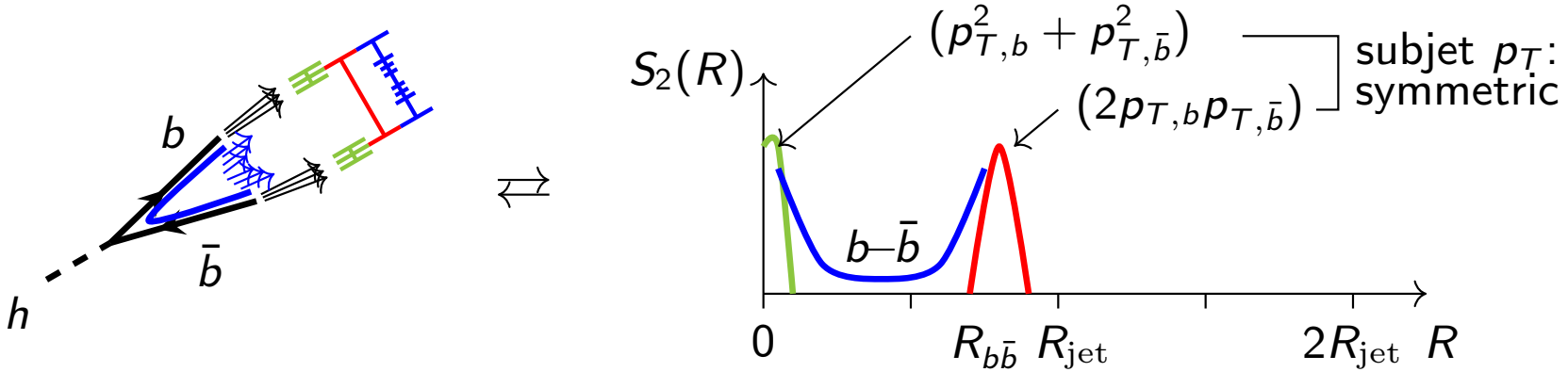
- The spectrum contains information on non-local correlation in the jet.



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

A typical Higgs jet

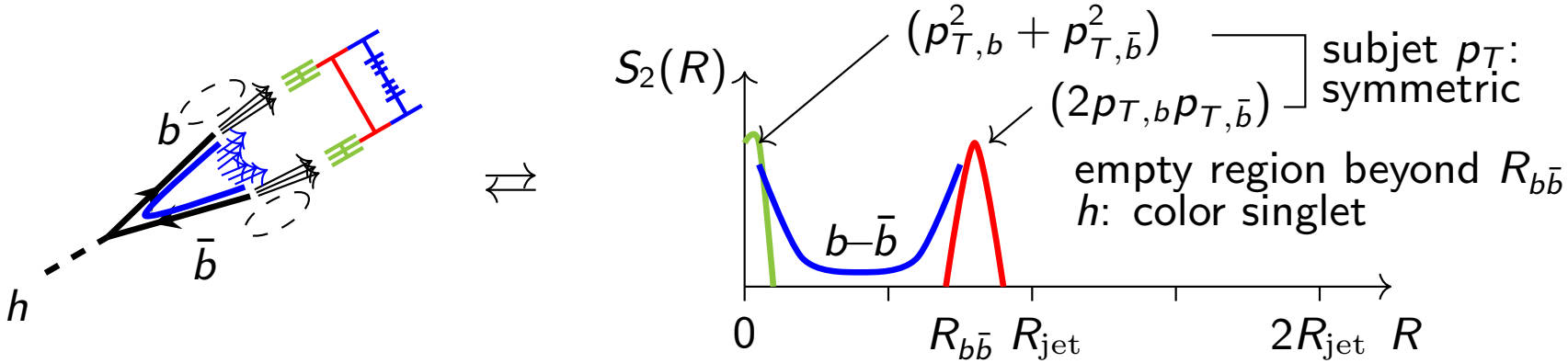
- The spectrum contains information on non-local correlation in the jet.



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

A typical Higgs jet

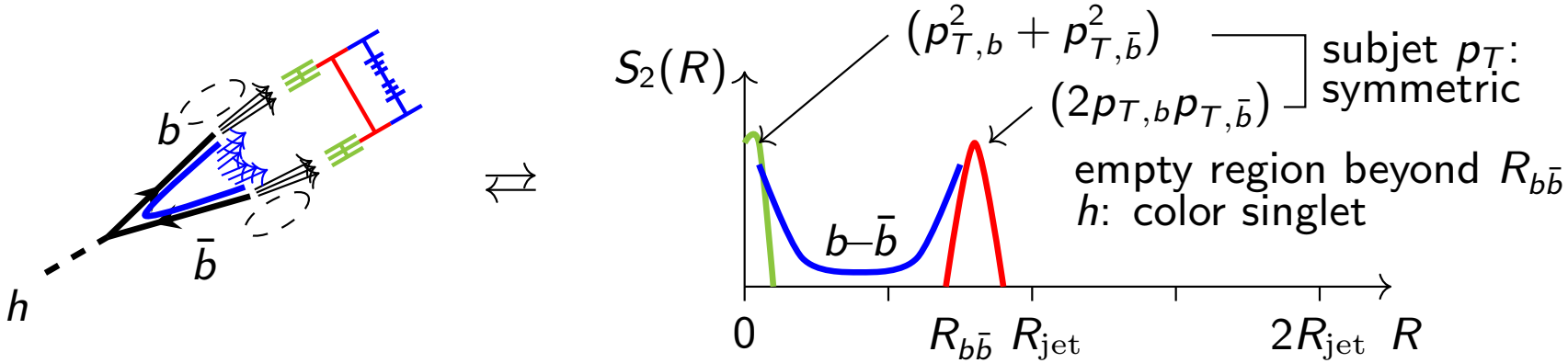
- The spectrum contains information on non-local correlation in the jet.



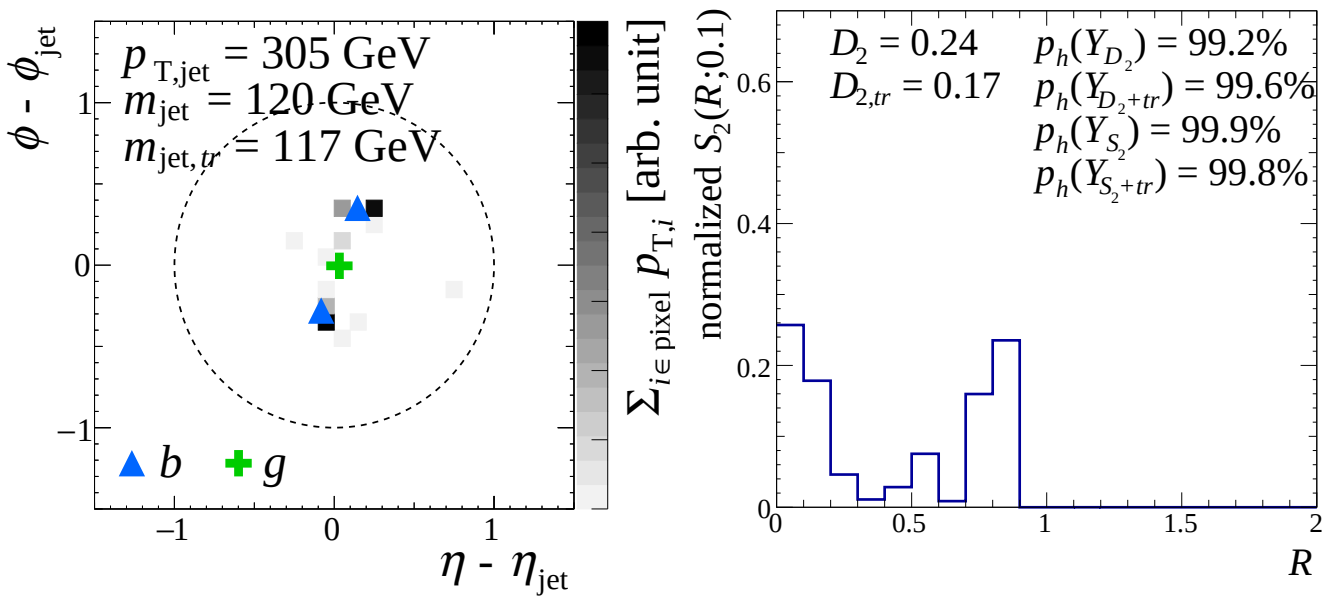
$$S_2(R;\Delta R) = \frac{1}{\Delta R} \sum_{\substack{i,j\in \text{jet} \\ R_{ij}\in[R,R+\Delta R)}} p_{T,i}p_{T,j}$$

A typical Higgs jet

- The spectrum contains information on non-local correlation in the jet.

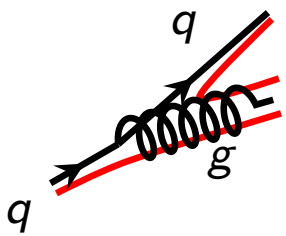


- This is a typical Higgs jet from MG5_aMC@NLO+Pythia8+Delphes.

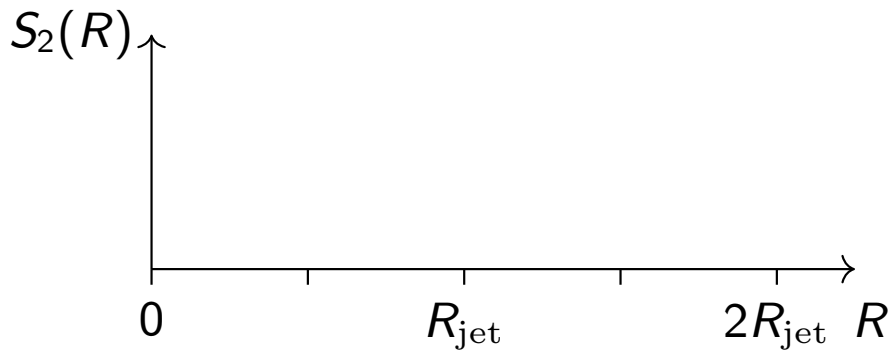


A typical QCD jet

- Meanwhile, a typical QCD jet has a smoothly falling behavior according to the parton splitting kernels.



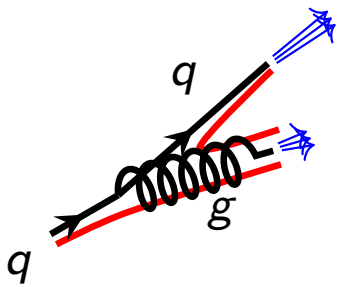
$\rightleftarrows$



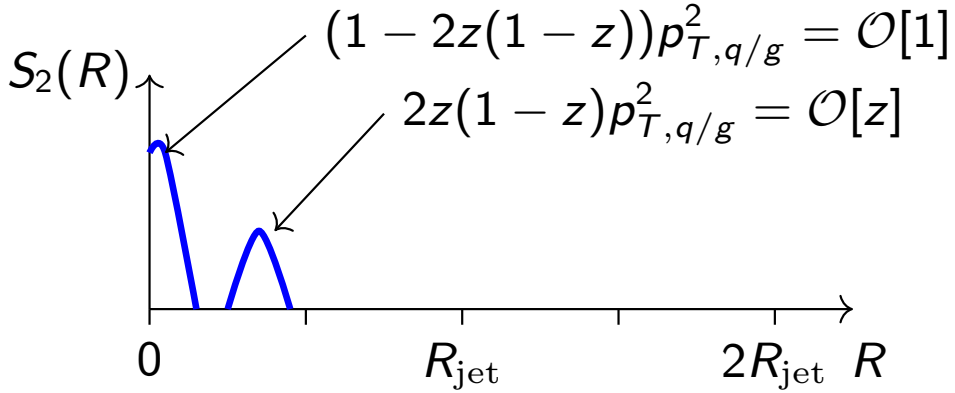
$$S_2(R;\Delta R) = \frac{1}{\Delta R} \sum_{\substack{i,j\in\text{jet} \\ R_{ij}\in[R,R+\Delta R)}} p_{T,i}\, p_{T,j}$$

A typical QCD jet

- Meanwhile, a typical QCD jet has a smoothly falling behavior according to the parton splitting kernels.



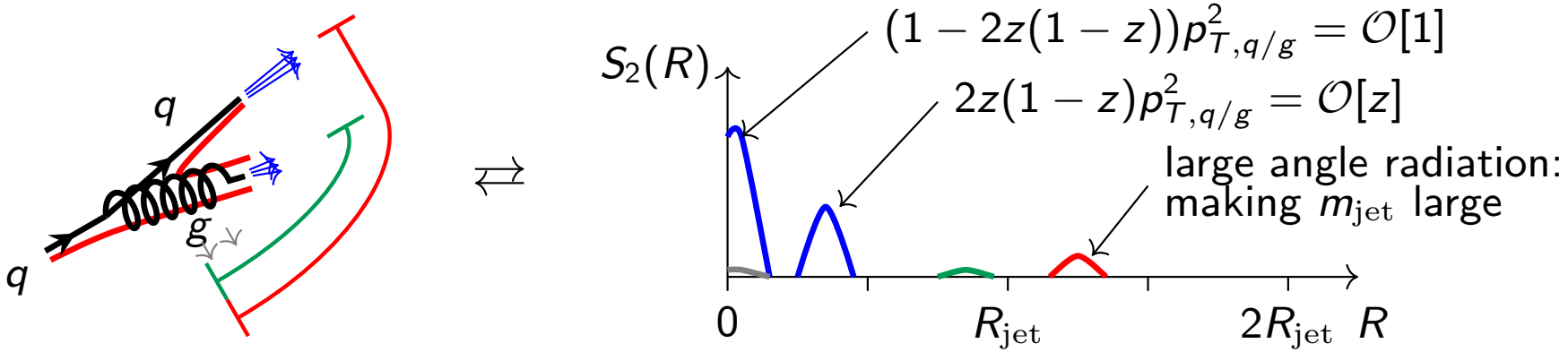
$\rightleftarrows$



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

A typical QCD jet

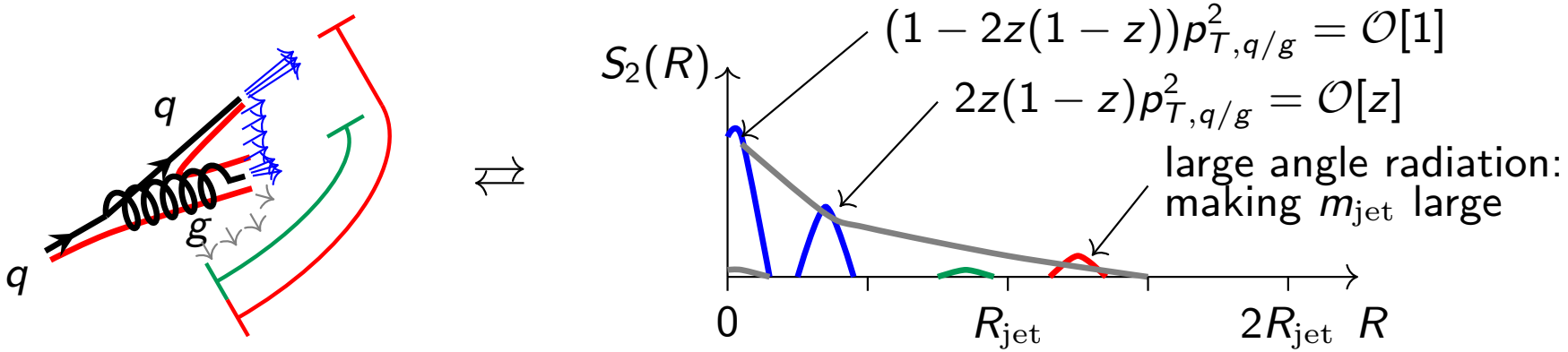
- Meanwhile, a typical QCD jet has a smoothly falling behavior according to the parton splitting kernels.



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

A typical QCD jet

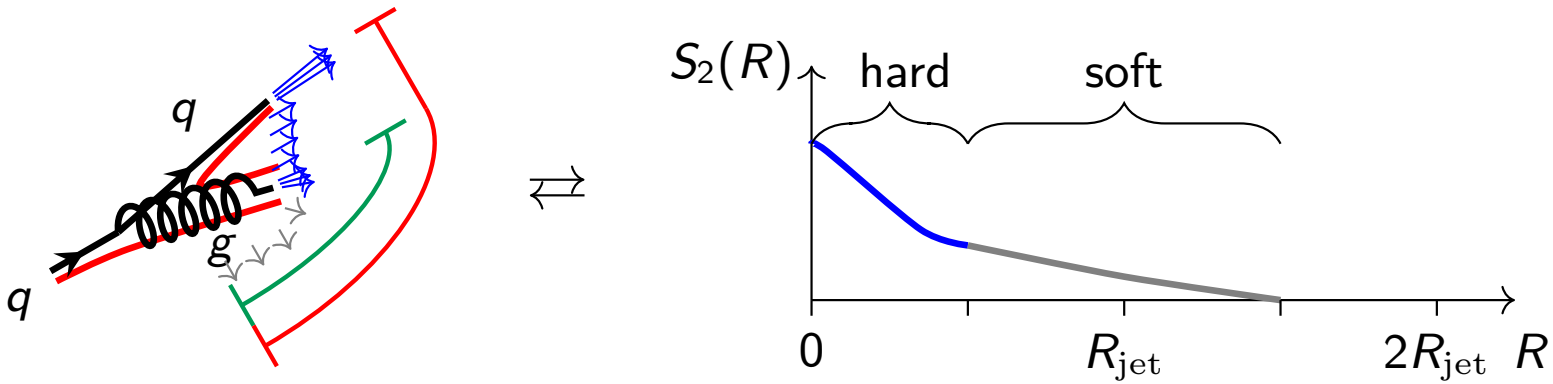
- Meanwhile, a typical QCD jet has a smoothly falling behavior according to the parton splitting kernels.



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

A typical QCD jet

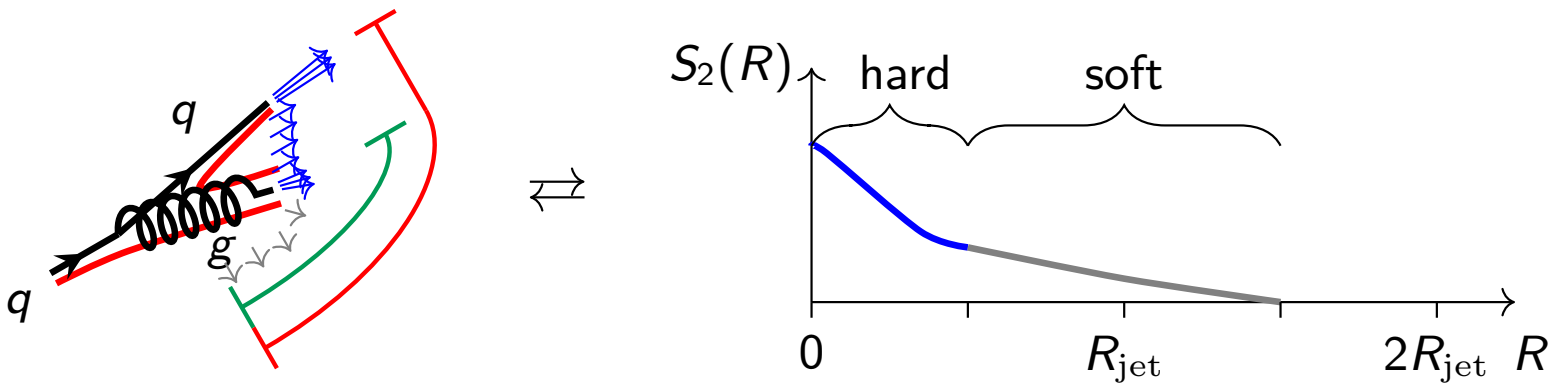
- Meanwhile, a typical QCD jet has a smoothly falling behavior according to the parton splitting kernels.



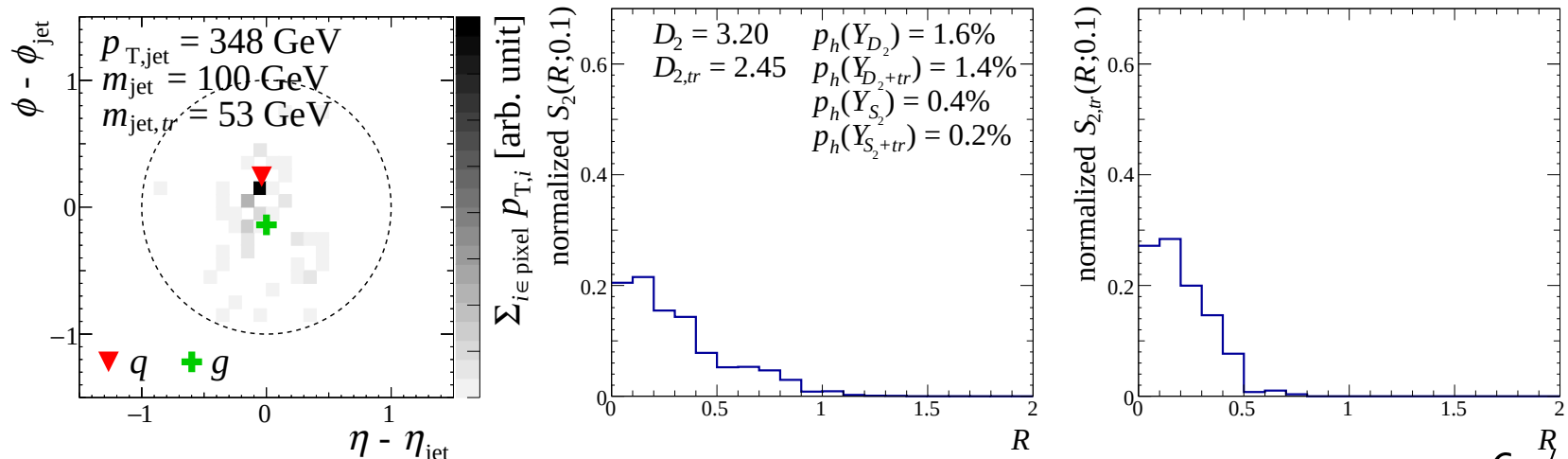
$$S_2(R;\Delta R) = \frac{1}{\Delta R} \sum_{\substack{i,j\in\text{jet} \\ R_{ij}\in[R,R+\Delta R)}} p_{T,i}\, p_{T,j}$$

A typical QCD jet

- Meanwhile, a typical QCD jet has a smoothly falling behavior according to the parton splitting kernels.



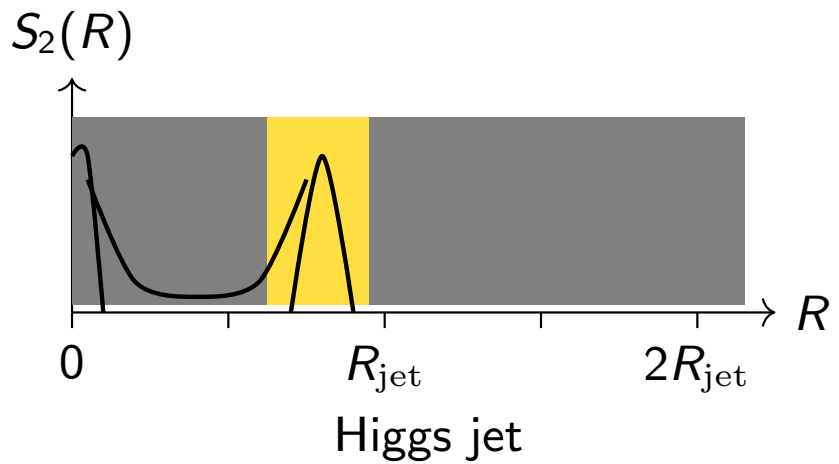
- This is a typical quark jet from MG5_aMC@NLO+Pythia8+Delphes.



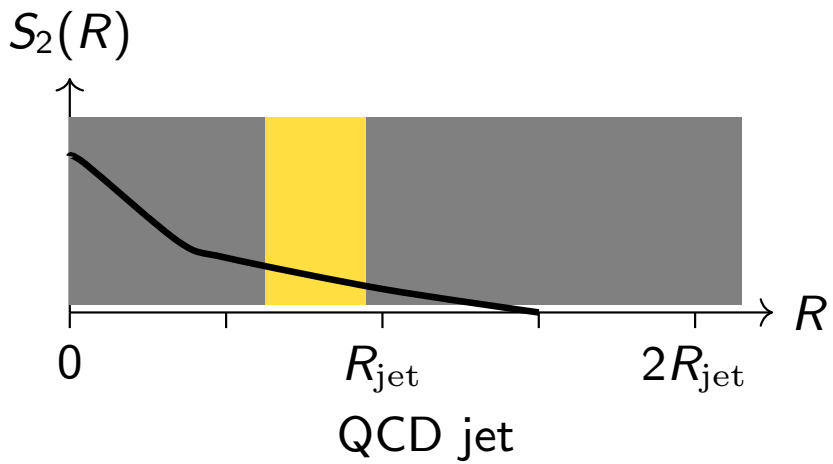
Intermediate Check: a quick classifier between the Higgs jets and QCD jets

- The typical Higgs jets and QCD jets have a different spectrum. The energy deposit difference could be used as a quick classifier.

$$R_{S_2} = \frac{\int_{a\hat{R}_{b\bar{b}}}^{\min[a'\hat{R}_{b\bar{b}}, R_{\rm jet}]} dR\, S_2(R)}{\int_0^{a\hat{R}_{b\bar{b}}} dR\, S_2(R) + \int_{\min[a'\hat{R}_{b\bar{b}}, R_{\rm jet}]}^{\infty} dR\, S_2(R)}, \qquad \hat{R}_{b\bar{b}} = \frac{2 \cdot 125\,\mathrm{GeV}}{p_{T,\rm jet}}$$



Higgs jet

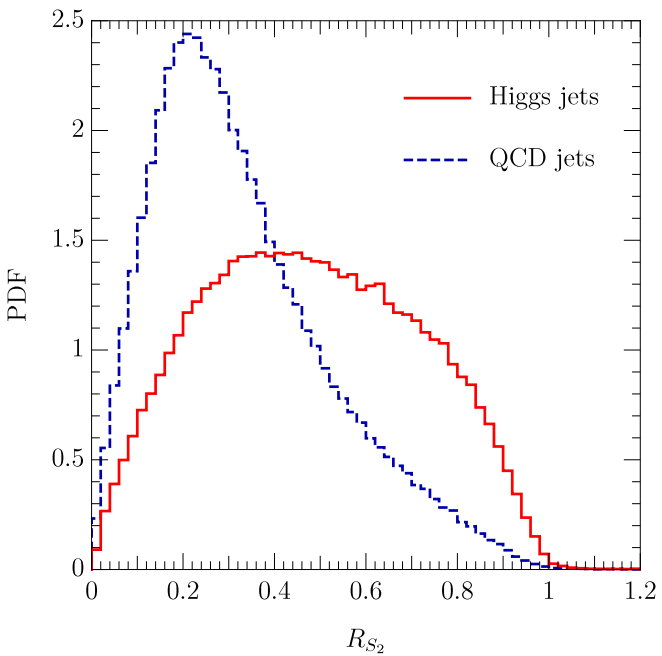


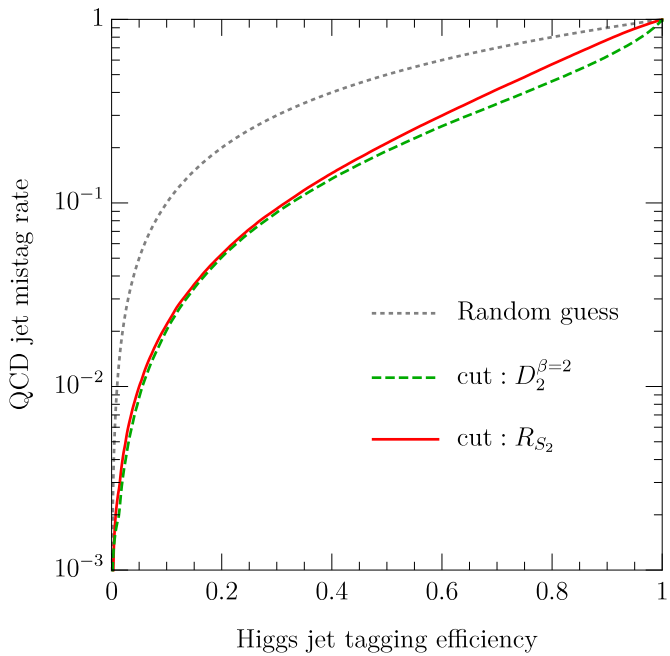
QCD jet

Intermediate Check: a quick classifier between the Higgs jets and QCD jets

- The typical Higgs jets and QCD jets have a different spectrum. The energy deposit difference could be used as a quick classifier.

$$R_{S_2} = \frac{\int_{a\hat{R}_{b\bar{b}}}^{\min[a'\hat{R}_{b\bar{b}}, R_{\rm jet}]} dR\, S_2(R)}{\int_0^{a\hat{R}_{b\bar{b}}} dR\, S_2(R) + \int_{\min[a'\hat{R}_{b\bar{b}}, R_{\rm jet}]}^{\infty} dR\, S_2(R)}, \qquad \hat{R}_{b\bar{b}} = \frac{2 \cdot 125\,\mathrm{GeV}}{p_{T,\rm jet}}$$





- Event preselection: $m_{\rm jet} \in [100, 150]$ GeV, and $p_{T,\rm jet} \in [300, 400]$ GeV.

An Artificial Neural Network?

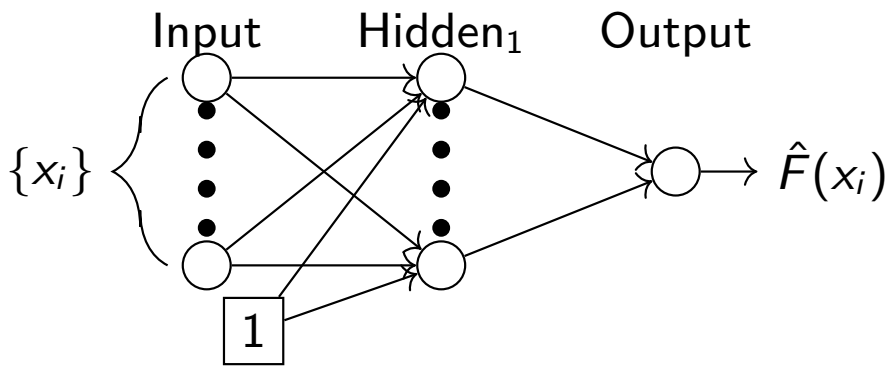
- Nevertheless, the energy deposit ratio of the spectrum is not the only information available in $S_2(R)$.
- We will introduce a neural network to use all the information available in $S_2(R)$.
- An artificial neural network is a mathematical model of a function.

$$\hat{F}(x_i) = v_j \cdot f(w_{ij} \cdot x_i + b_j)$$

where f is a nonlinear function, for example, we may consider ReLU,

$$f_{\mathrm{ReLU}}(x) = \begin{cases} x & x > 0 \\ 0 & x <= 0 \end{cases} \tag{1}$$

A graphical representation is follows.



- "Universal approximation theorem" states that if we have an enough number of the hidden neurons, which is the hidden dimension of j, then we can approximate any bounded continuous functions.

Our ANN setup

- Hence, we trained a neural network for the Higgs jet and QCD jet classification, and used the network as a classifier.

$$\begin{aligned} z_i^{(0)} &= x_i \\ z_i^{(n)} &= f_{\mathrm{ReLU}}\left(w_{ij}^{(n)} z^{(n-1)} + b_i^{(n)}\right) \\ \hat{y}_i &= \exp\left[w_{ij}^{(5)} z_j^{(4)}\right] / \sum_i \exp\left[w_{ij}^{(5)} z_j^{(4)}\right] \end{aligned}$$

- This network is trained for predicting a class vector

$$\vec{y} = \begin{cases} (1,0) & \text{Higgs jet} \\ (0,1) & \text{QCD jet} \end{cases}$$

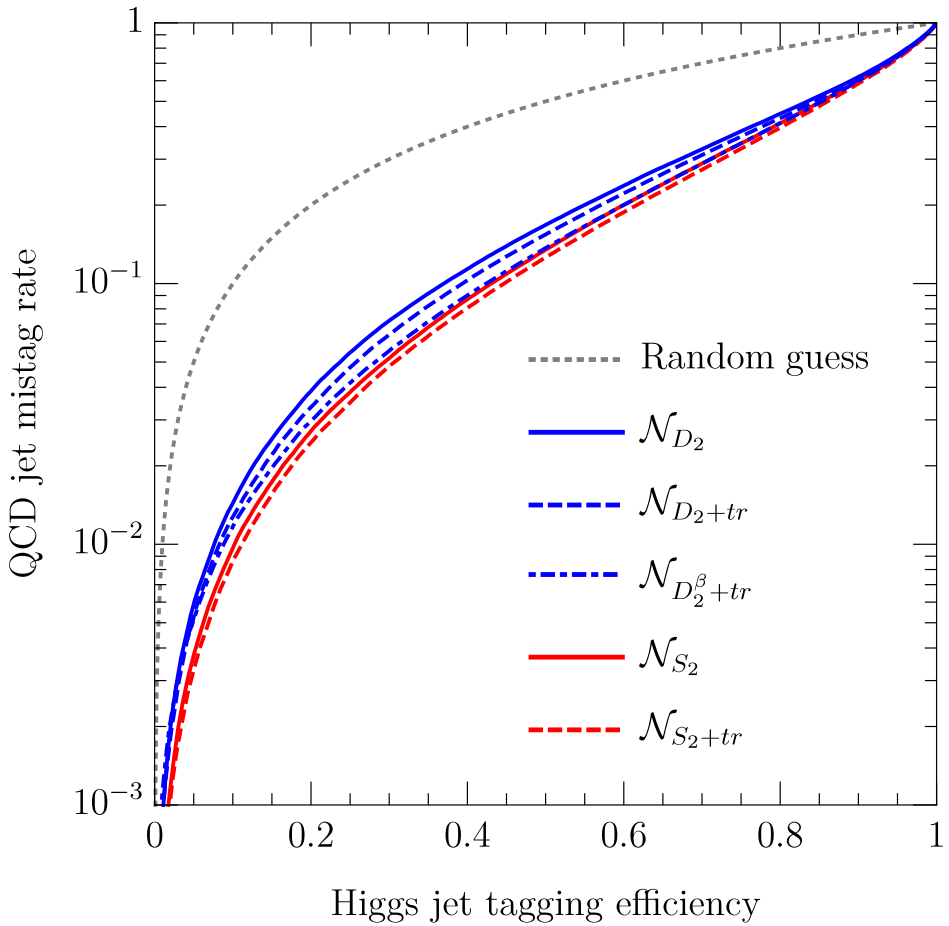
by minimizing the categorical cross entropy

$$L = -\sum_{\text{event}} \sum_i y_i \log \hat{y}_i$$

- The predicted class vector $\hat{y}$, which is the output of the neural network, can be used as a classifier.

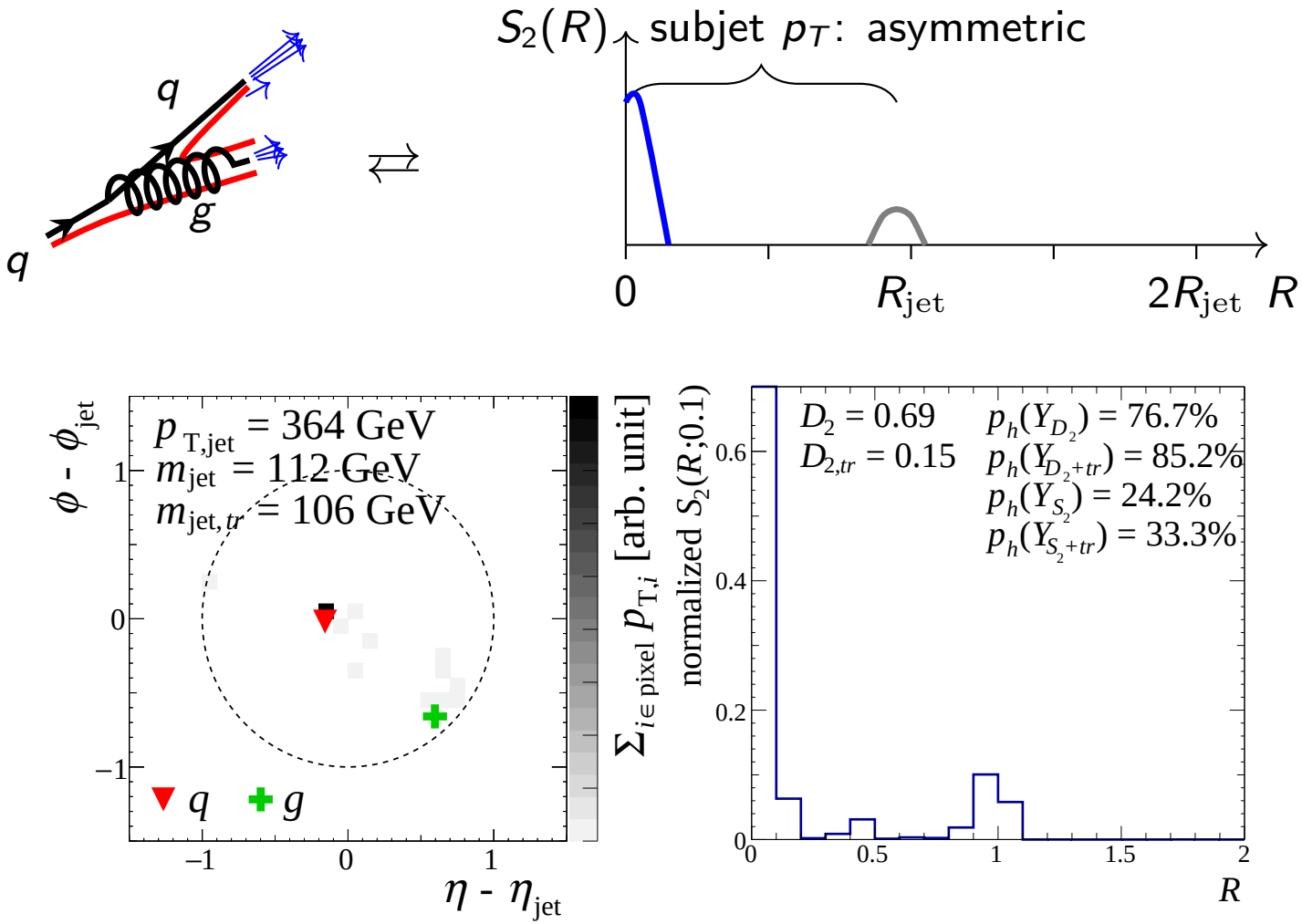
Comparing $S_2(R)$ and D_2

- To compare the information contained in $S_2(R)$ and D_2, we compare the ROC curves of neural networks with folowing inputs.
- $\mathcal{N}_{D_2}$: $\{p_{T,\mathrm{jet}}, m_{\mathrm{jet}}, D_2^{\beta=2}\}$
- $\mathcal{N}_{S_2}$: $\{p_{T,\mathrm{jet}}, m_{\mathrm{jet}}, S_2(0;0.1), \cdots, S_2(1.9;0.1)\}$
- $+tr$: trimmed variables are included.
- Q: what's the source of the difference?
  - Reduce mistagging
  - Going beyond two-prong



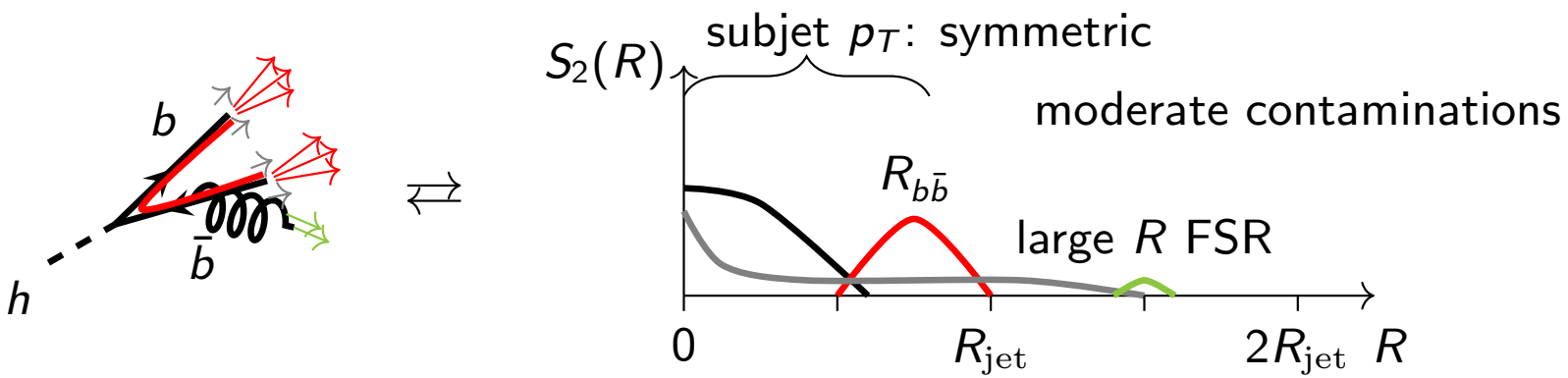
- Event preselection: $m_{\mathrm{jet}} \in [100, 150]$ GeV, and $p_{T,\mathrm{jet}} \in [300, 400]$ GeV.

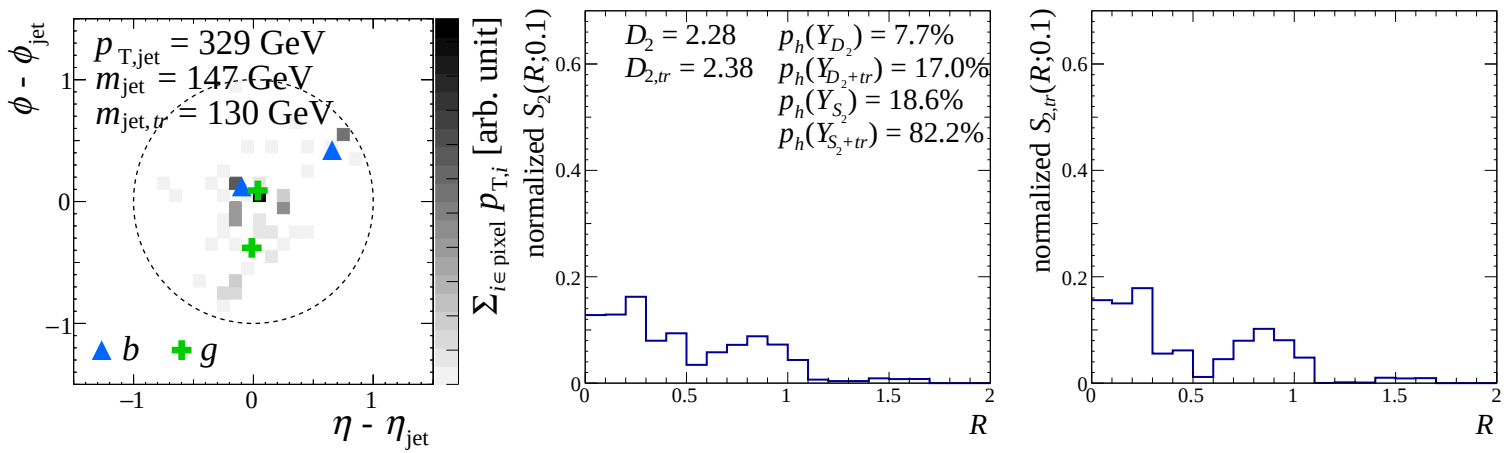
A jet tagged in D_2 analysis only



- $\mathcal{N}_{D_2}$: $\{p_{T,\rm jet}, m_{\rm jet}, D_2^{\beta=2}\}$
- Even though this jet is two-prong, the likelihood for QCD jet is not that small.

A jet tagged in S_2 analysis with trimming only





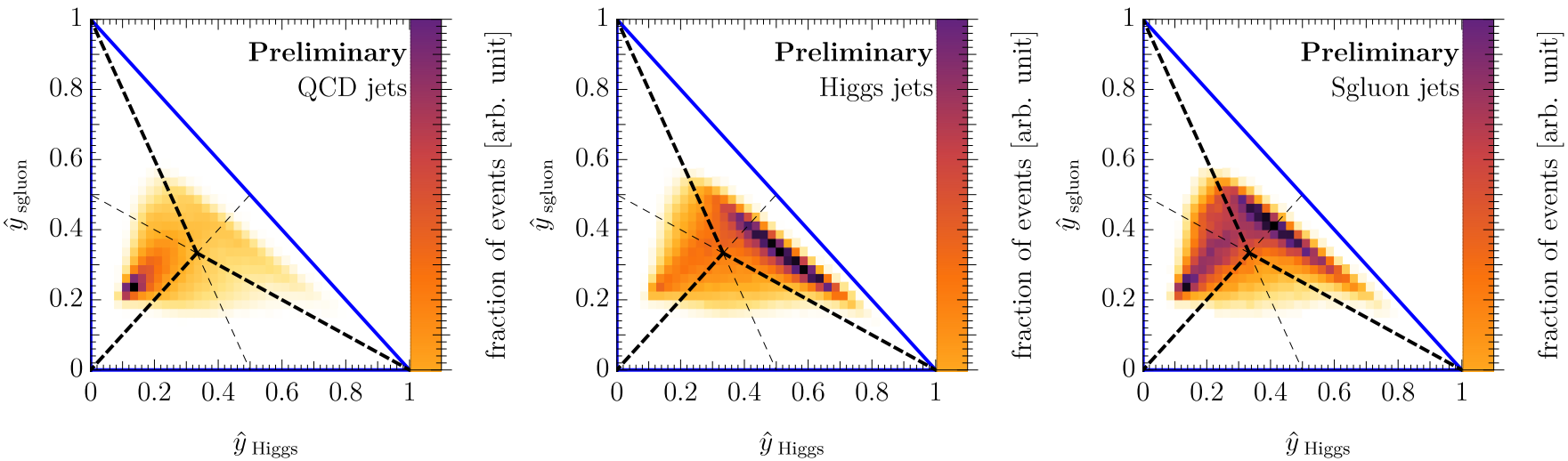
- This jet has two-prong substructure, but the subjets are wide and contaminated by other QCD activities.
- $S_2(R)$ together with jet trimming and a neural network can identify hard and soft substructure.

Color charge of the boosted Higgs jets

Preliminary

- $S_2(R)$ shows a good performance on identifying boosted Higgs jets from QCD jets.
- Is it possible to identify color charge of the Higgs jet from $S_2(R)$?
- We considered a color octet scalar, sgluon, decaying to $b\bar{b}$ and did a multi-class classification with a neural network $\hat{y}$ predicting

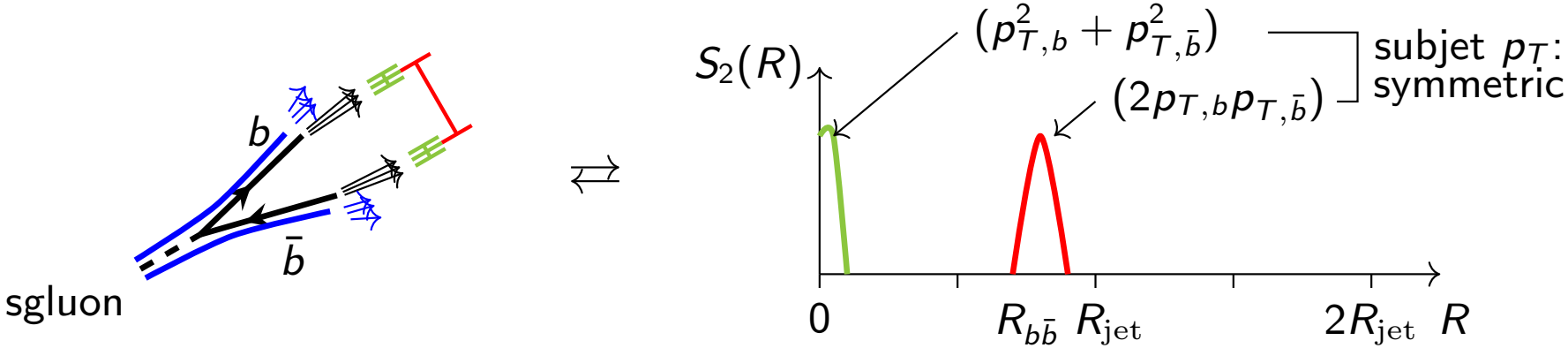
$$\vec{y} = \begin{cases} (1,0,0) & \text{Higgs jet} \\ (0,1,0) & \text{Sgluon jet} \\ (0,0,1) & \text{QCD jet} \end{cases}$$



A typical Sgluon jet

Preliminary

- The sgluon jet has a different radiation pattern compared to the Higgs jet because of its nontrivial color charge.

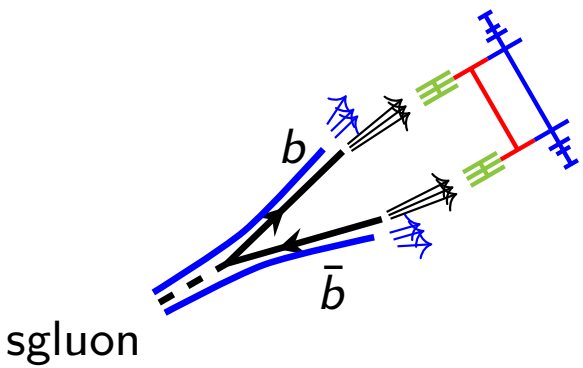


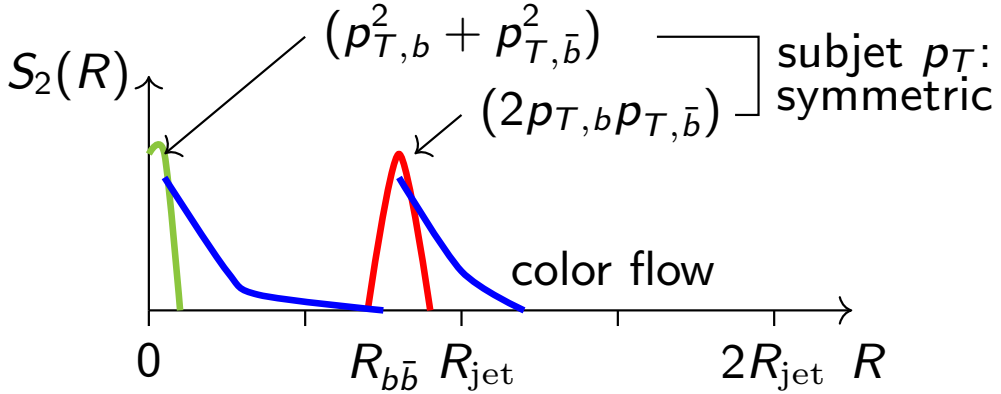
$$S_2(R;\Delta R) = \frac{1}{\Delta R} \sum_{\substack{i,j\in\text{jet} \\ R_{ij}\in[R,R+\Delta R)}} p_{T,i}p_{T,j}$$

A typical Sgluon jet

Preliminary

- The sgluon jet has a different radiation pattern compared to the Higgs jet because of its nontrivial color charge.



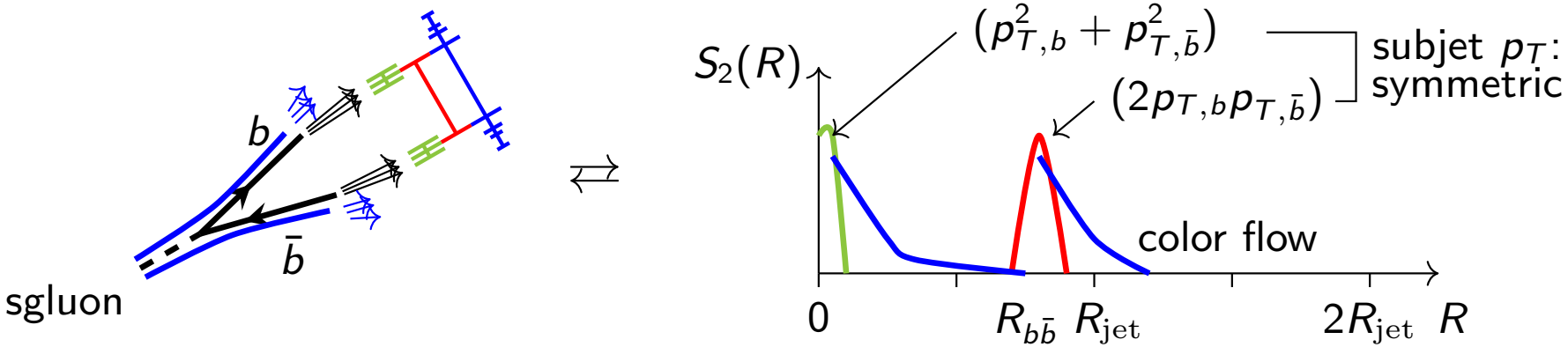


$$S_2(R;\Delta R) = \frac{1}{\Delta R} \sum_{\substack{i,j\in\text{jet} \\ R_{ij}\in[R,R+\Delta R)}} p_{T,i}p_{T,j}$$

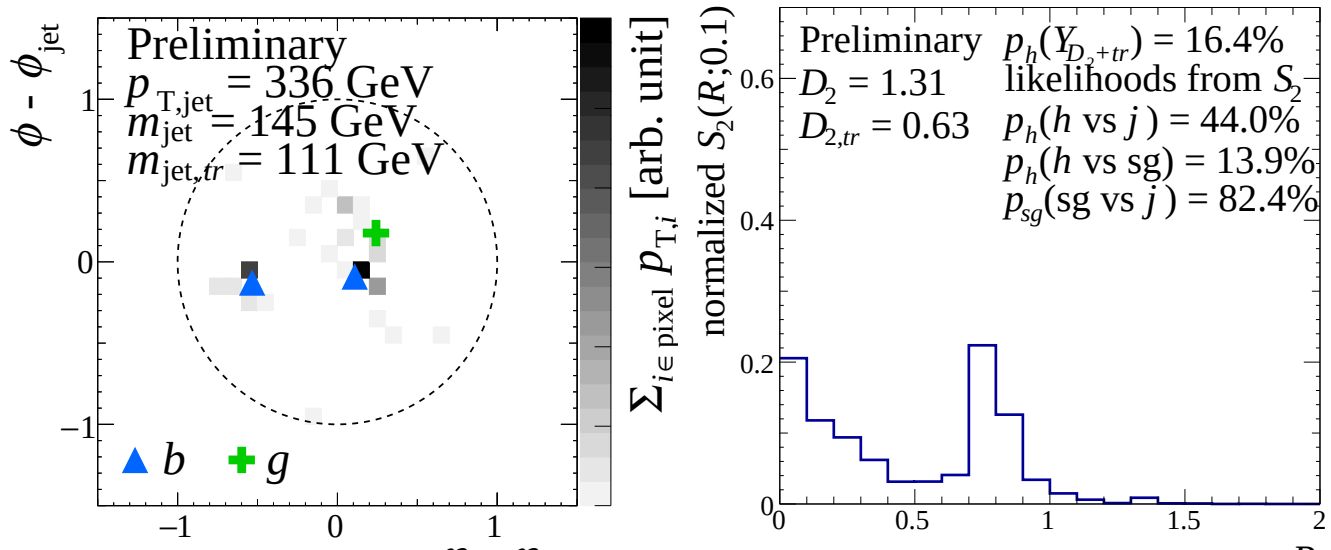
A typical Sgluon jet

Preliminary

- The sgluon jet has a different radiation pattern compared to the Higgs jet because of its nontrivial color charge.



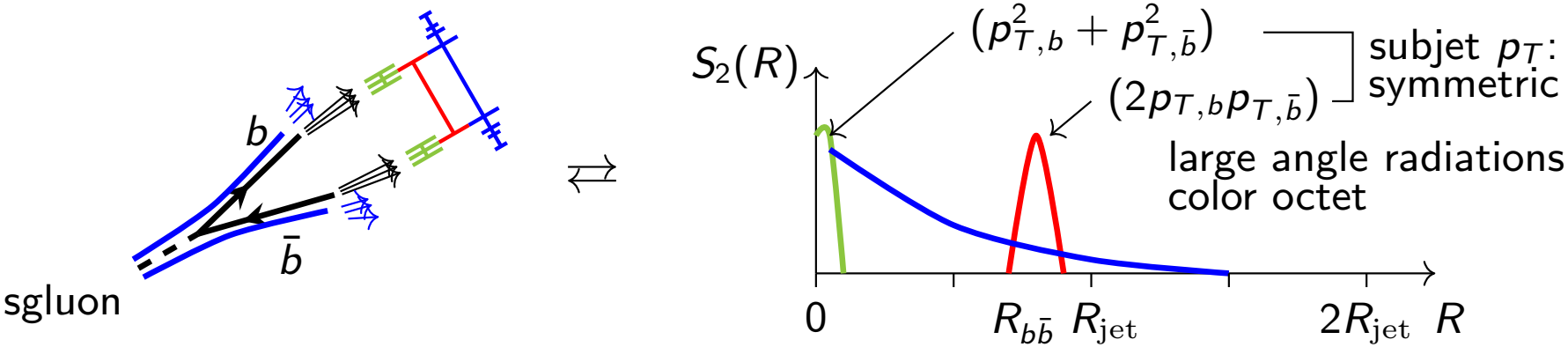
- This is a typical sgluon jet from MG5_aMC@NLO+Pythia8+Delphes.



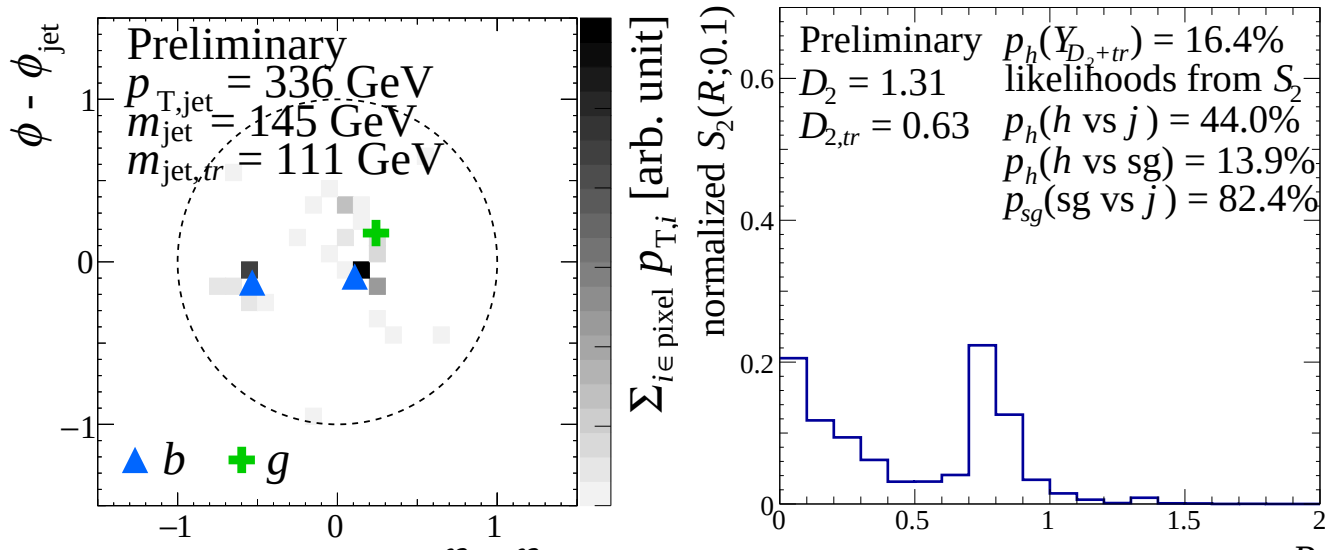
A typical Sgluon jet

Preliminary

- The sgluon jet has a different radiation pattern compared to the Higgs jet because of its nontrivial color charge.



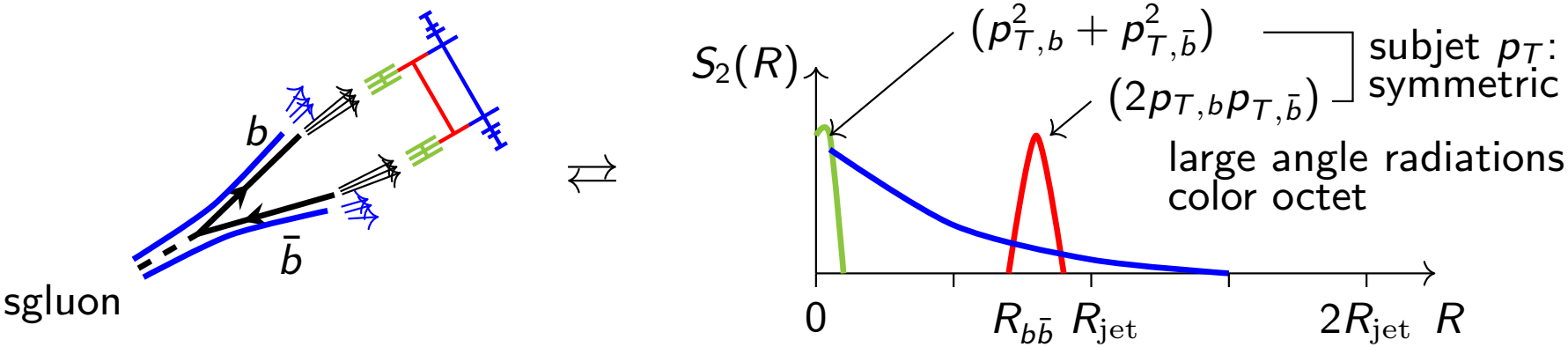
- This is a typical sgluon jet from MG5_aMC@NLO+Pythia8+Delphes.



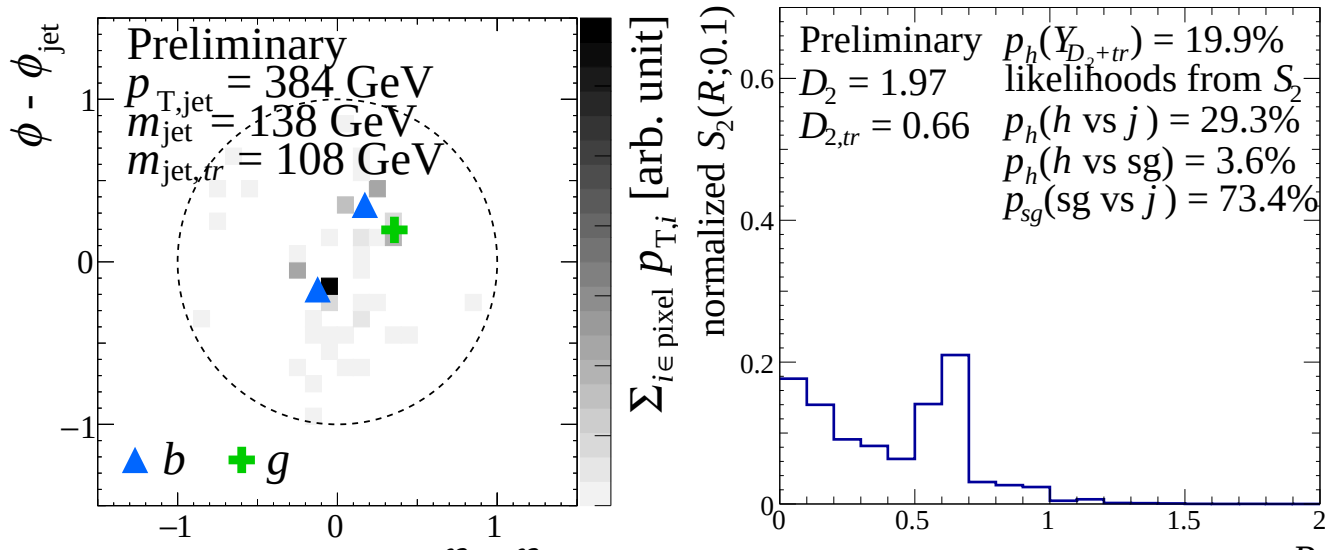
A typical Sgluon jet

Preliminary

- The sgluon jet has a different radiation pattern compared to the Higgs jet because of its nontrivial color charge.

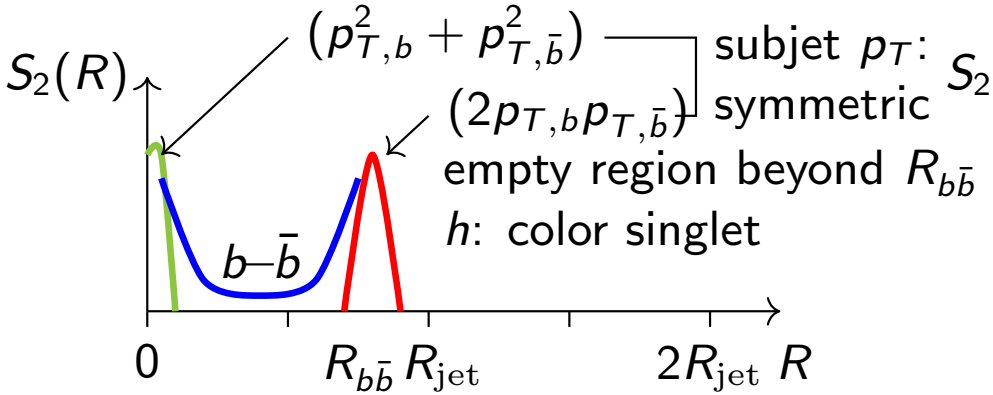


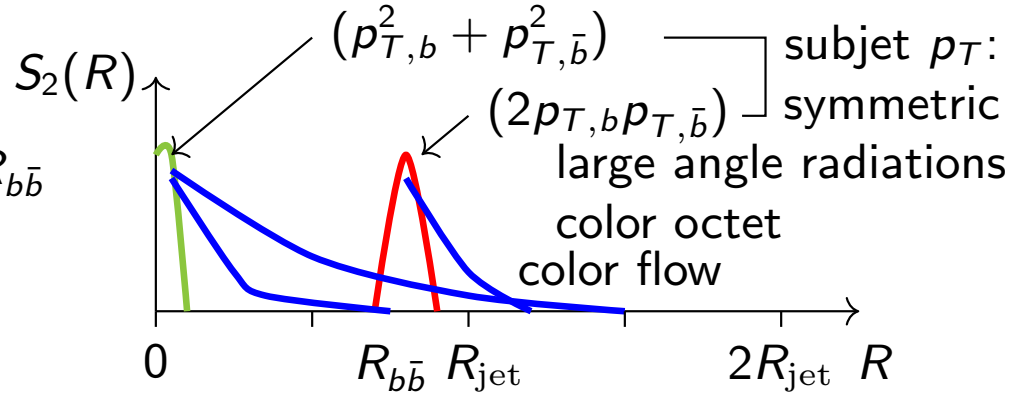
- This is a typical sgluon jet from MG5_aMC@NLO+Pythia8+Delphes.



Conclusion

- We have introduced a spectral analysis of jet substructure for classifying Higgs jets and QCD jets.
- The spetrum $S_2(R)$ is highly visual and useful in describing jet substructure with large angular scale.
- The improvement of performance of Higgs jet vs QCD jet classification is not large, but the spectral function $S_2(R)$ provide us a systematical framework for investigating jet substructure over various angular scale.
- For example, $S_2(R)$ can be used for understanding a color charge of boosted jets.



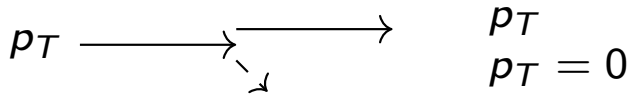


- There are more wider applications will be come out soon, so please stay tuned!

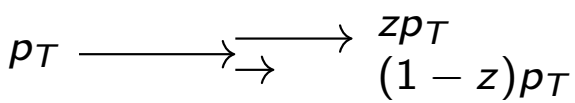
Backups

IRC safety of $S_2(R)$

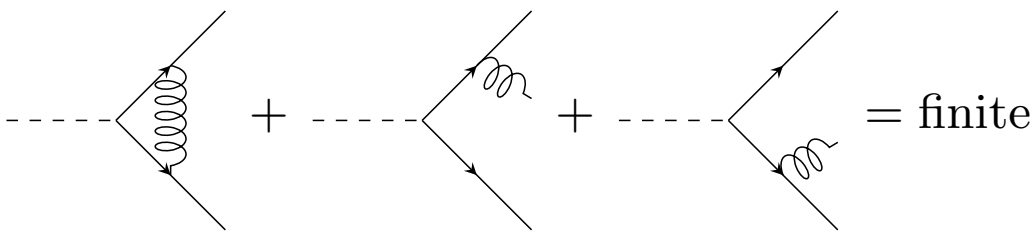
- The $S_2(R)$ spectrum is better to be infrared and collinear (IRC) safe, namely invariant under soft and collinear radiations.
  - Soft radiation: the parton radiates a $p_T = 0$ parton.



  - 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)



(2)

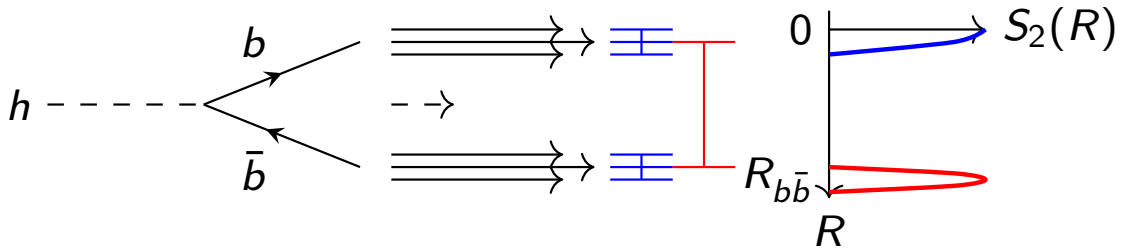
- $S_2(R)$ is IRC safe because soft or collinear radiation does not introduce any new angular scale R.

IRC safety of $S_2(R)$

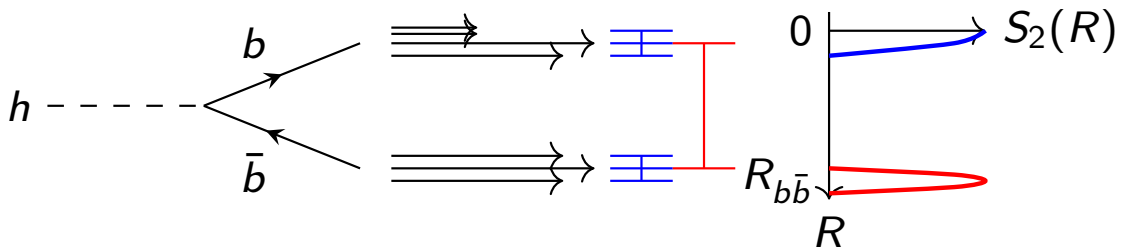
- The $S_2(R)$ spectrum is IRC safe.

$$S_2(R;\Delta R) = \frac{1}{\Delta R} \sum_{\substack{i,j\in \mathrm{jet} \\ R_{ij}\in[R,R+\Delta R)}} p_{T,i}p_{T,j}, \tag{3}$$

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



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



Comparison to another variable

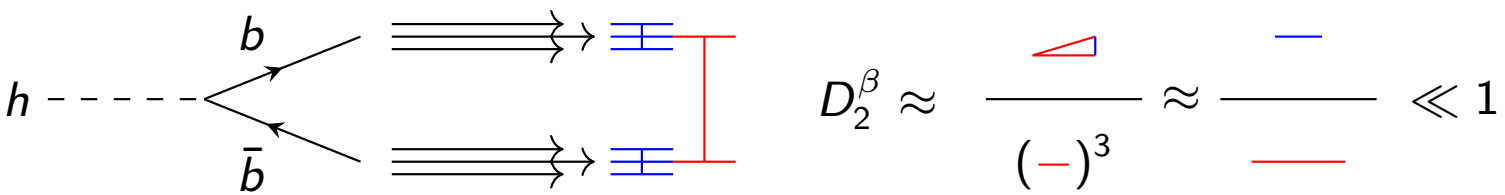
- For comparison, we prepare an ANN anlysis with D_2 variable (1409.6298)

$$e_2^\beta = \frac{1}{p_{T,\mathrm{jet}}^2} \sum_{\substack{i,j\in\mathrm{jet} \\ i<j}} p_{T,i} p_{T,j} R_{ij}^\beta, \tag{4}$$

$$e_3^\beta = \frac{1}{p_{T,\mathrm{jet}}^3} \sum_{\substack{i,j,k\in\mathrm{jet} \\ i<j<k}} p_{T,i} p_{T,j} p_{T,k} R_{ij}^\beta R_{jk}^\beta R_{ki}^\beta, \tag{5}$$

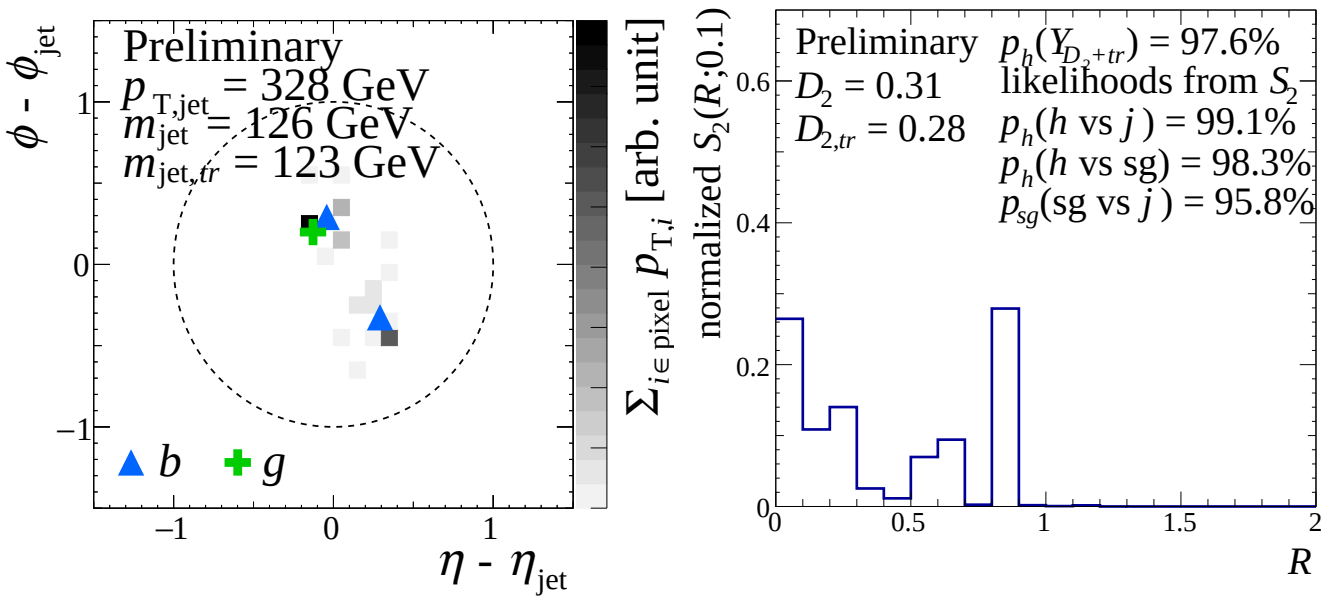
$$D_2^\beta = \frac{e_3^\beta}{(e_2^\beta)^3} \sim \frac{\triangle}{(-)^3}, \tag{6}$$

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



Higgs jet and color octet classifier

Preliminary



sgluon!

Preliminary