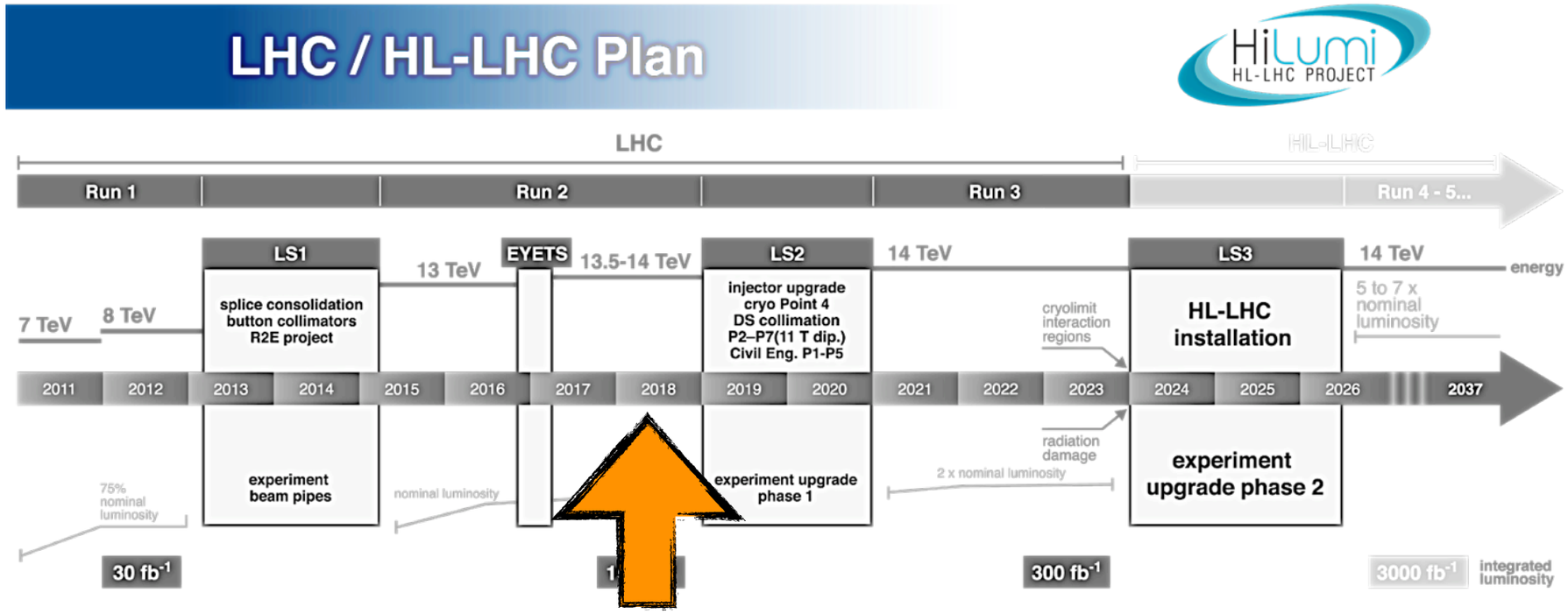
Quark jet rates and quark gluon discrimination in multi-jet final states

Yasuhito Sakaki
(KAIST)

1807.01421

ICHEP 2018 @Seoul, 4-11/7/2018



- No clear sign of BSM at the LHC
- Need to examine final states more precisely
- Final states are categorized by inclusive variables

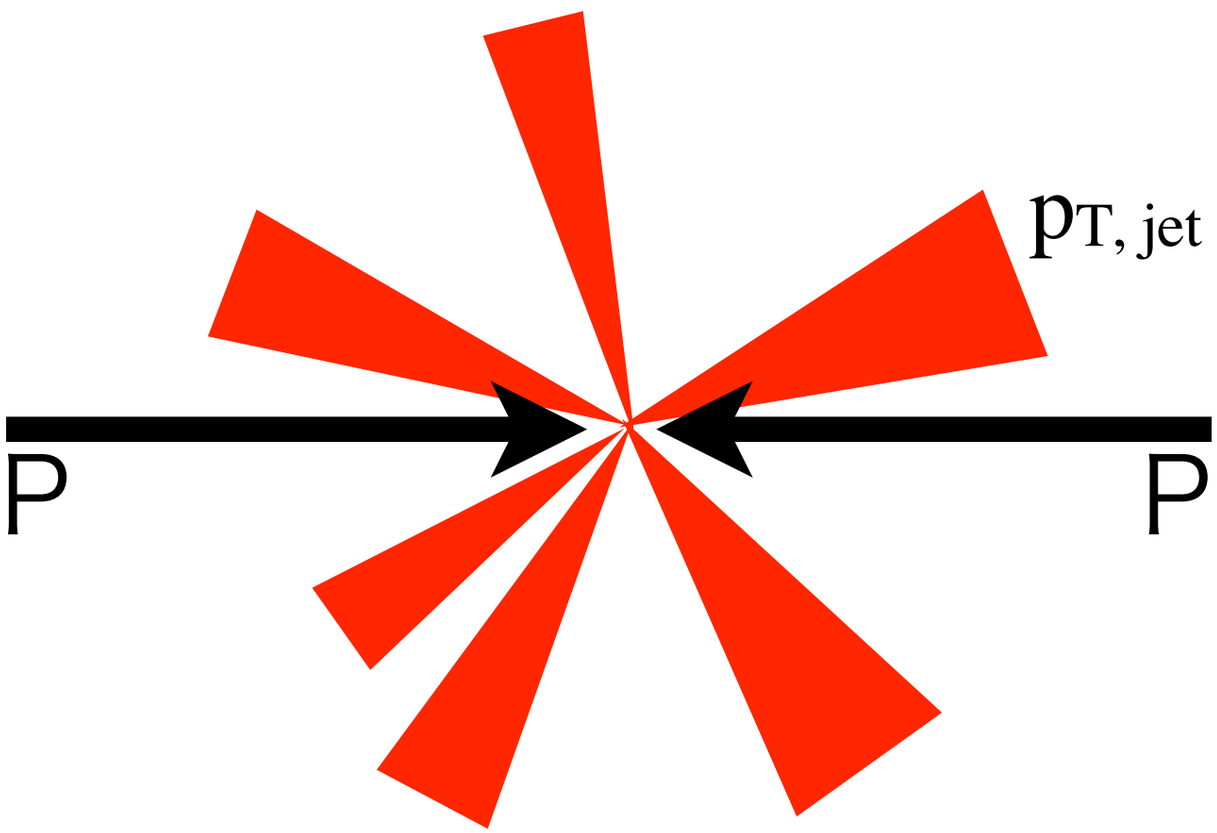
  ➡ N(jets), N(leptons), MET, H_T, ...

- Categories containing jets encounter a huge QCD background
- As increasing N(jets), kinematics and MC validation become more complicated
- LHC is jets production machine. We want to examine precisely even such multi-jet final state.

Multi-jet final state and New physics

- **No accurate simulation** for the large jet multiplicity background due to the absence of higher-order, huge number of diagrams...

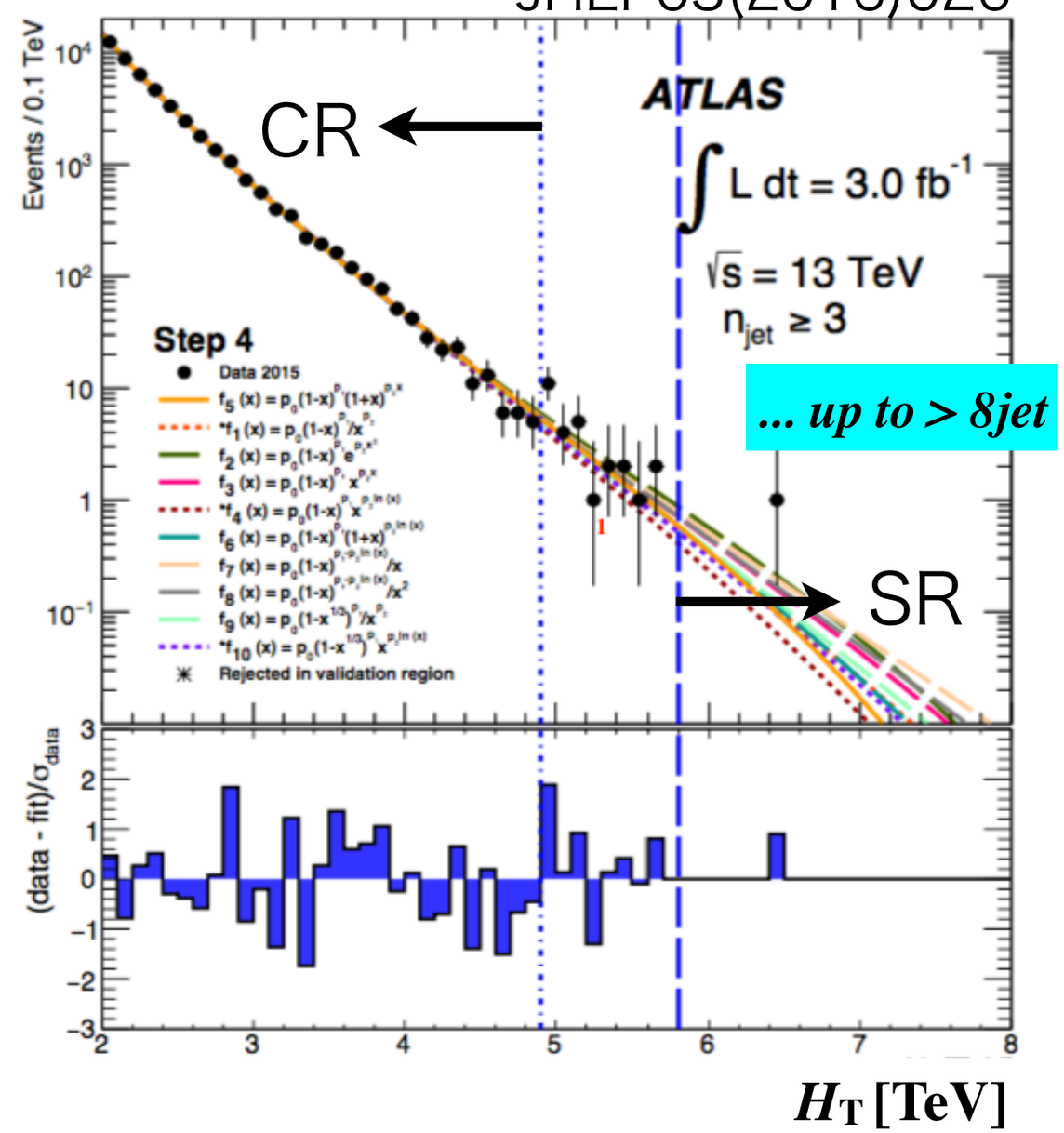
  ➡ **data-driven analysis**



inclusive variable:

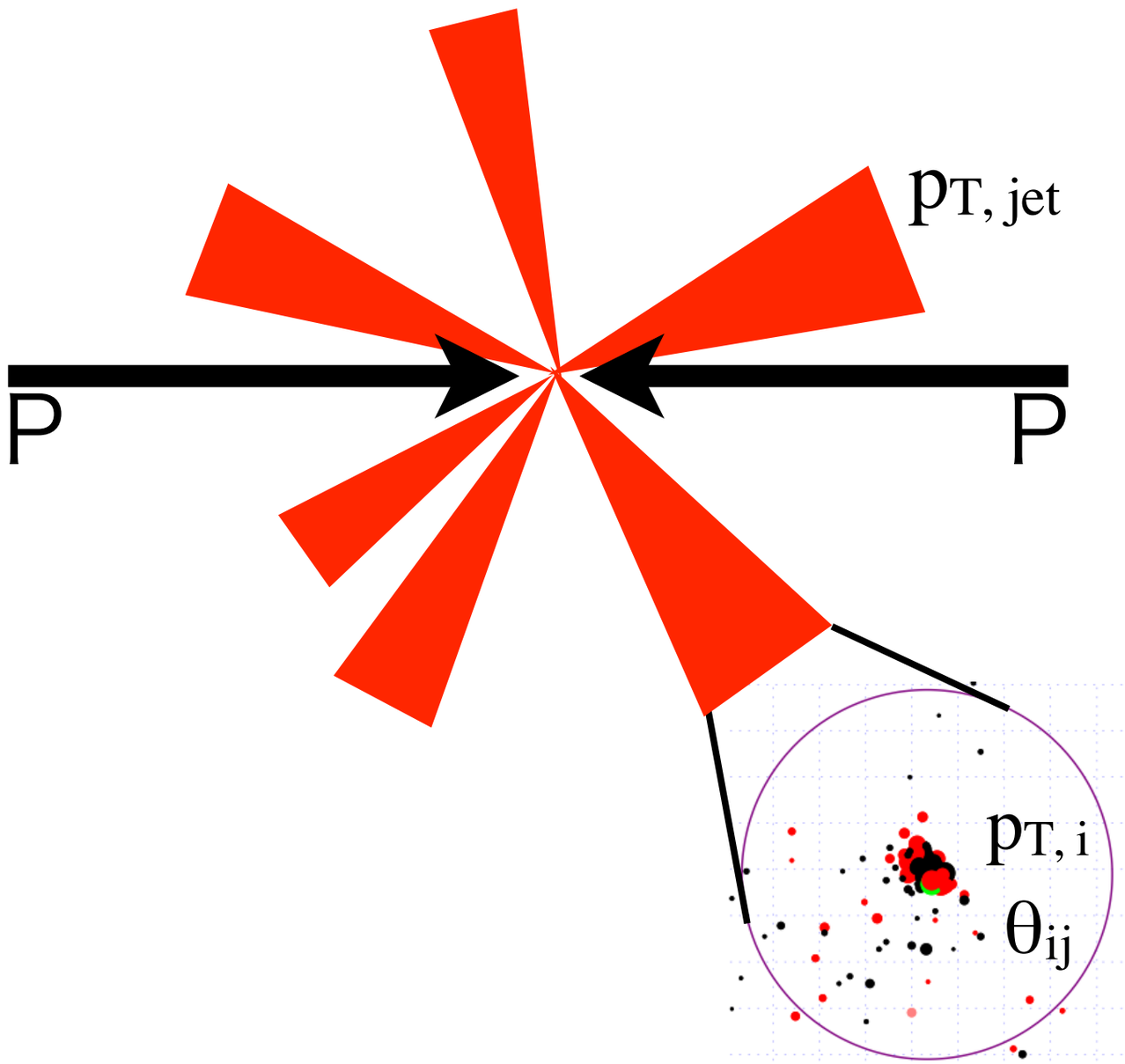
ex) $H_T = \Sigma p_{T,\,jet}$

JHEP03(2016)026



Multi-jet final state and New physics

- Accurate simulation for the large jet multiplicity background does not exist due to the absence of higher-order, huge number of diagrams...



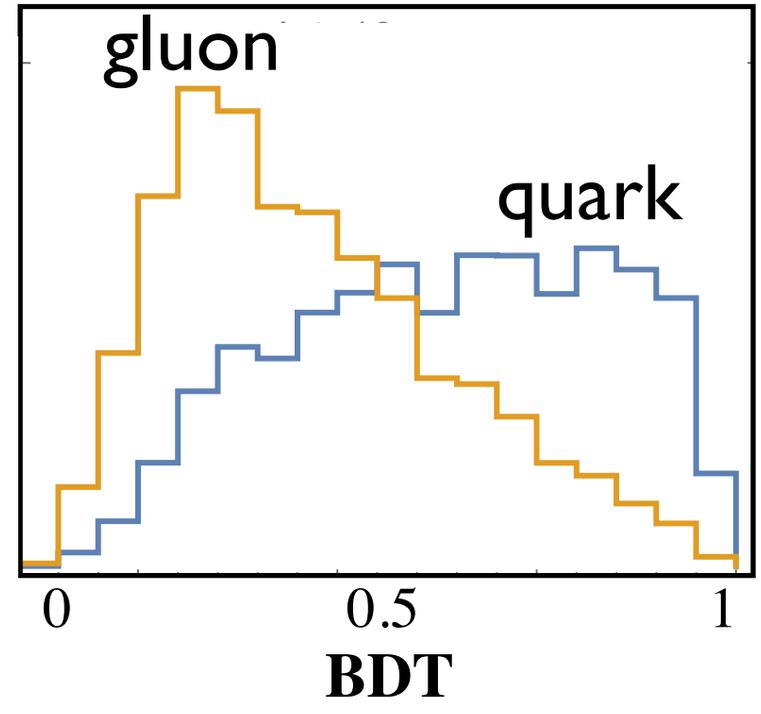
inclusive
$\Sigma p_{T,\,jet}$

exclusive
$\{p_{T,\,jet}\}$

jet flavor/substructure
$\{p_{T,\,i}\},\{\theta_{ij}\}$

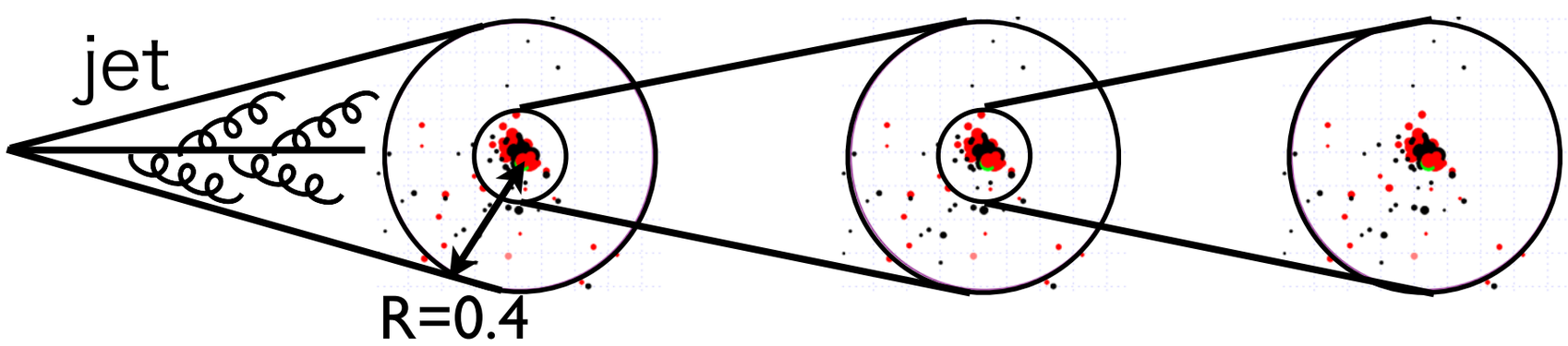
Jet substructure techniques

∋ Quark/Gluon discrimination

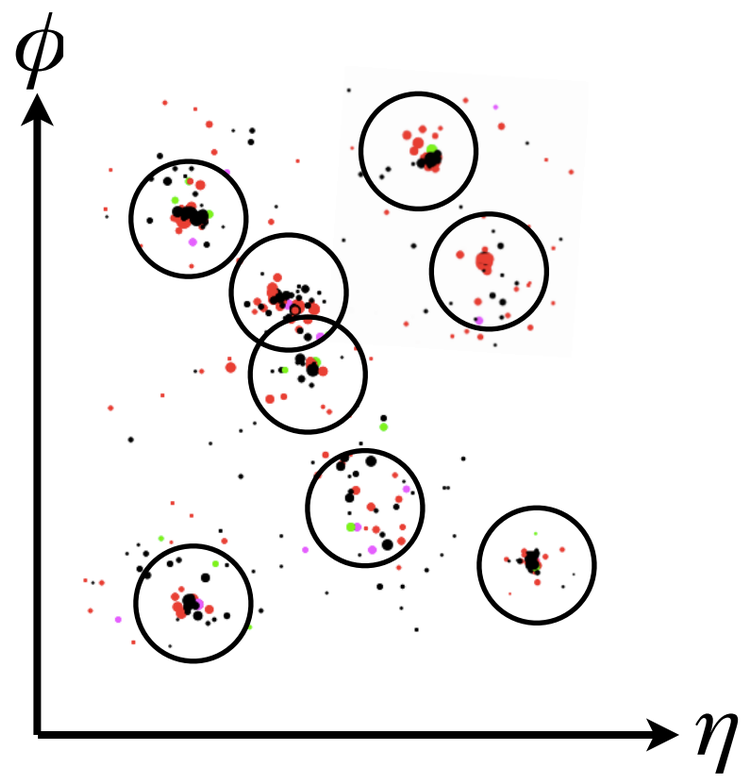


- It works well even with small-R (even with R<0.4)

- QCD radiation is approximately scale invariant



Multi-jet final state and Jet substructure



$$\Delta N = \left| N_{\text{quark}}^{\text{signal}} - N_{\text{quark}}^{\text{BKG}} \right|$$

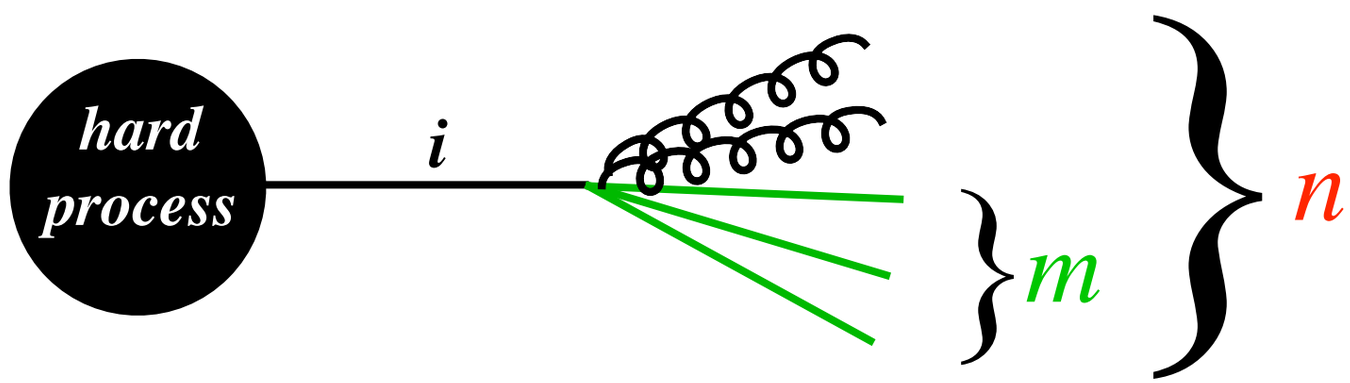
$$\Rightarrow \frac{S}{B} \propto \left(\frac{\epsilon_{\text{quark}}}{\epsilon_{\text{gluon}}} \right)^{\Delta N}, \qquad \frac{\epsilon_{\text{quark}}}{\epsilon_{\text{gluon}}} > 1$$

- Large enhancement of S/B for a signal that predicts the number of quark jets which is different from what the QCD background does

➡ Let's see **how many quark jets are included in the QCD background**.

Quark jet rates

- $R^i_{n,m}$: Probability that i emits n jets in which m quark jets are contained



- Generating functional:

N(jets) N(quark jets)

$$\Phi_i(p,t) = \sum_{n=1}^{\infty} \sum_{m=0}^{n} u^n v^m R^i_{n,m}(p,t)$$

$$\Phi_i(p,t) = u\, v_i\, \Delta_i(p,t) + \sum_k \int_{p_0/p}^{1} dz \int_{t_0}^{t} \frac{dt'}{t'} \frac{\Delta_i(p,t)}{\Delta_i(p,t')} \mathcal{P}_{i\to jk}\, \Phi_j(p,t')\, \Phi_k(zp,t')$$

Solutions

N(jets) N(quark jets)

$$\Phi_i(p,t) = \sum_{n=1}^{\infty} \sum_{m=0}^{n} u^n v^m R^i_{n,m}(p,t)$$

$$\Phi_q = u^1 v^1 \, e^{-(1-u)a_q L^2} \, e^{S_q} \cdots$$

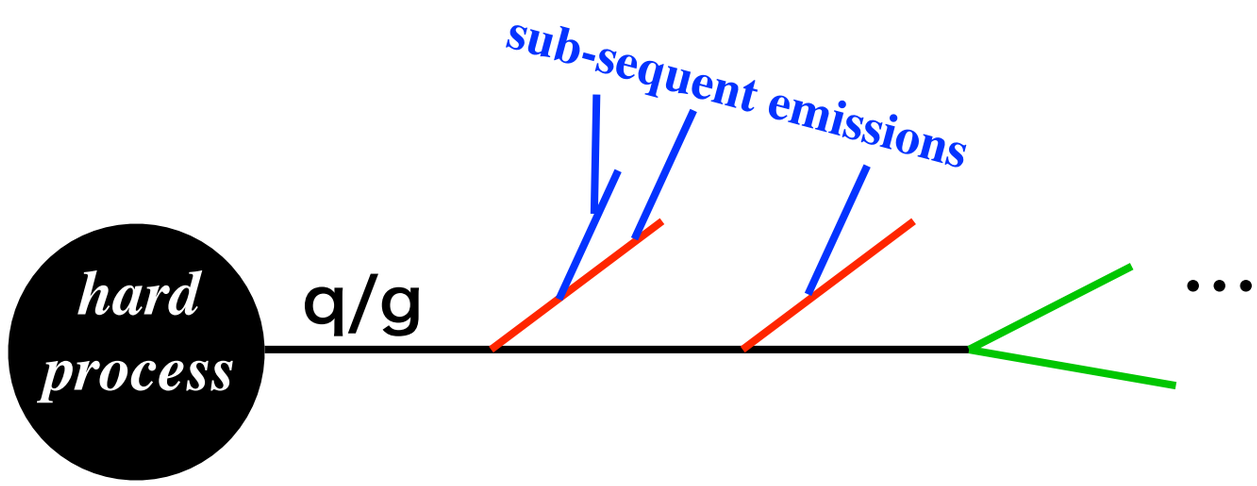
$$\Phi_g = u^1 v^0 \, e^{-(1-u)a_g L^2} \, e^{-(1-uv^2)a_{qq} L} \, e^{S_g} \cdots$$

0-th order

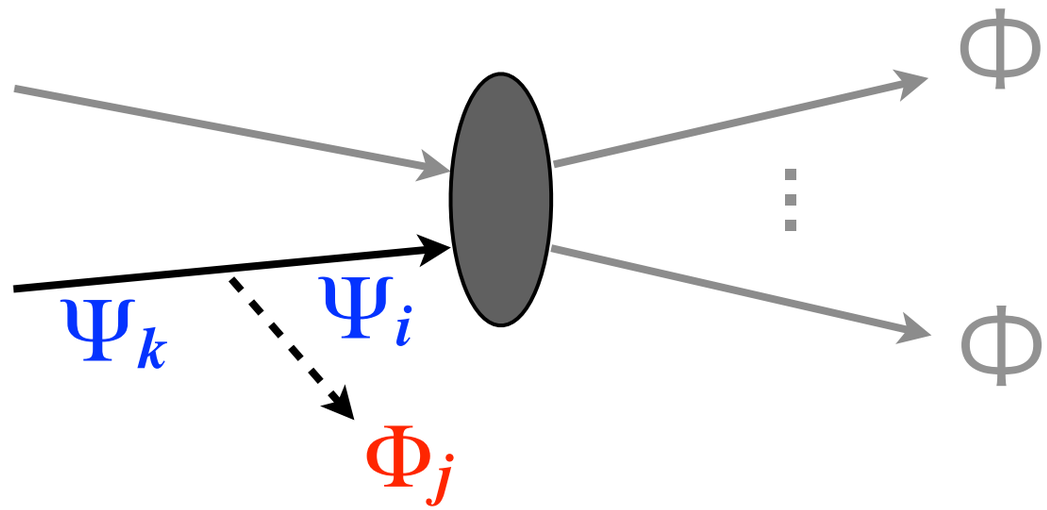
DL (gluon jet)

SL(g→qq) (quark jet)

subsequent emissions (both)



ISR



- evolution equation

$$\Psi_i(x,t) = \Pi_i(x,t) + \sum_k \int_x^1 \frac{dx'}{x'} \int_{t_0}^t \frac{dt'}{t'} \frac{\Pi_i(p,t)}{\Pi_i(p,t')} \times \frac{f_k(x',t)}{f_i(x,t)} \mathcal{P}_{k\to ij}\, \Psi_k(x',t')\, \Phi_j((x'-x)\,p_b, t')$$

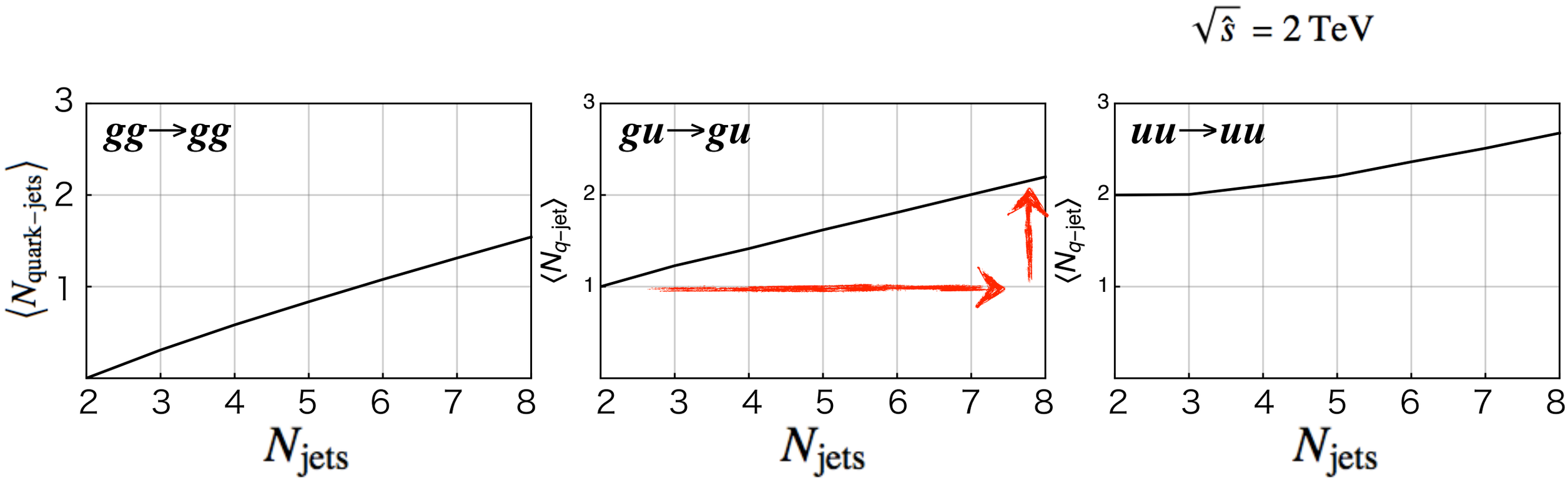
- solutions

$$\Psi_q = e^{-(1-u)\,a_q\,\kappa_q\,\lambda}\; e^{-(1-uv)\,c_q\,a_{qq}\,\lambda}\; e^{S_g[f_{q/q}]} \;\cdots$$

$$\Psi_g = e^{-(1-u)\,a_g\,\kappa_g\,\lambda}\; e^{-(1-uv)\,c_g\,a_q\,\lambda}\; e^{S_g[f_{g/g}]} \;\cdots$$

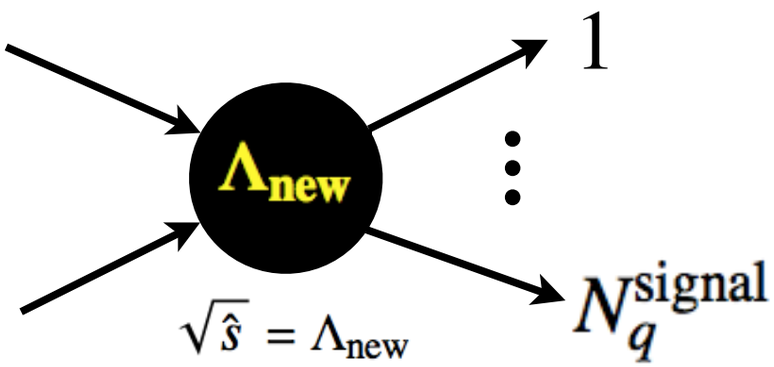
(gluon jet) (quark jet) subsequent emissions (both)

of quark jets



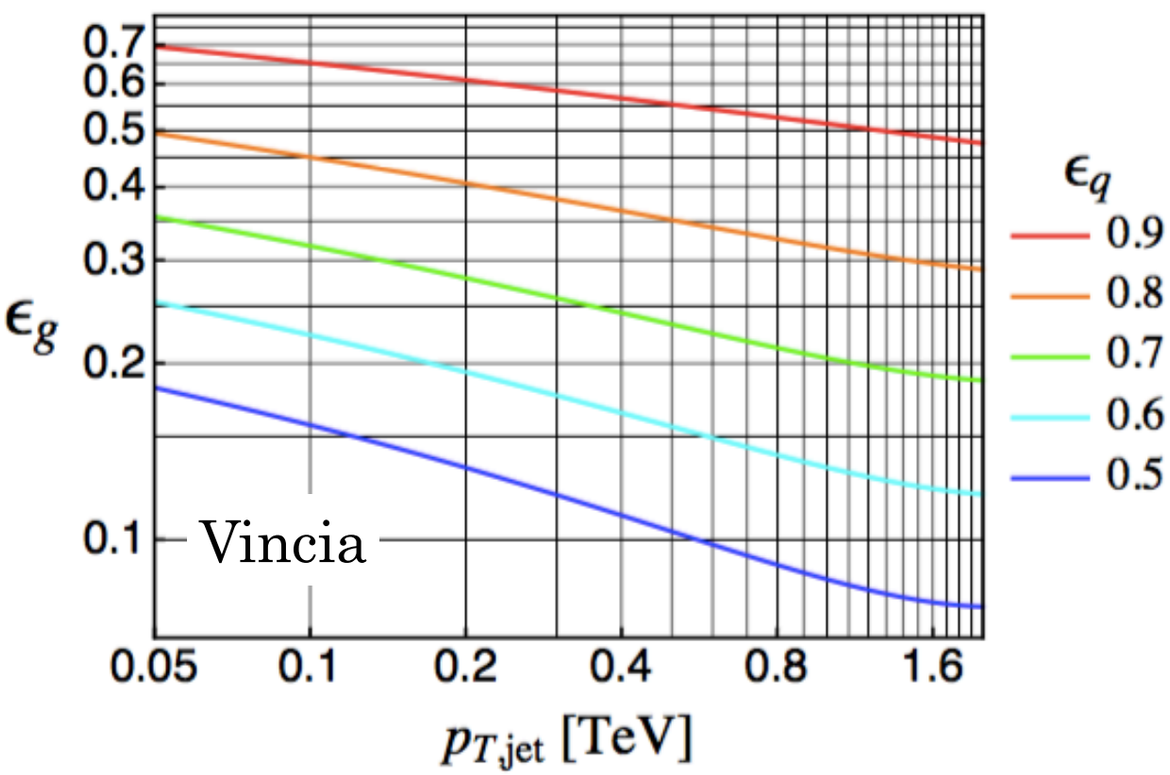
- Increase of gluon jet (double-log), quark jet (single-log)
- QCD multi-jets background is composed of few valence quark jets and many gluon jets
- W/Z/gamma + jets are also available
- It would be useful for MC tuning and development

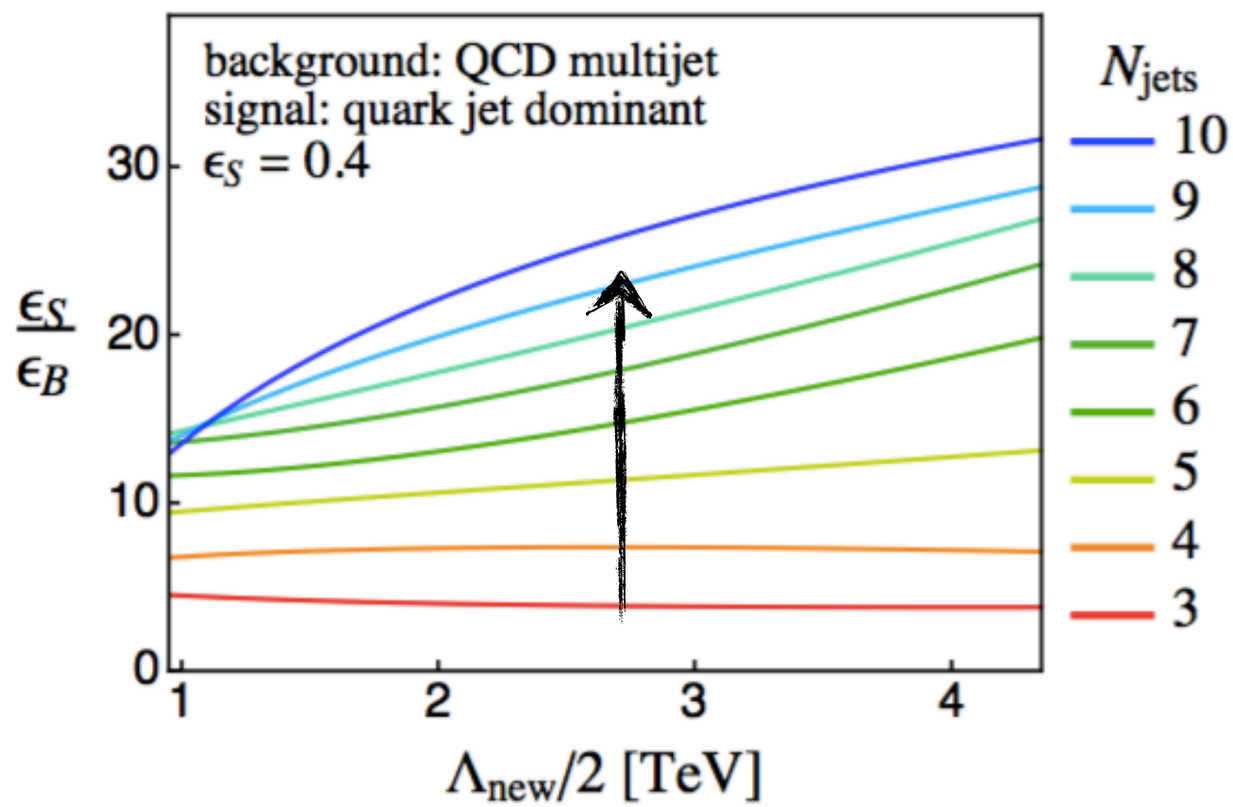
Expected improvement of S/B



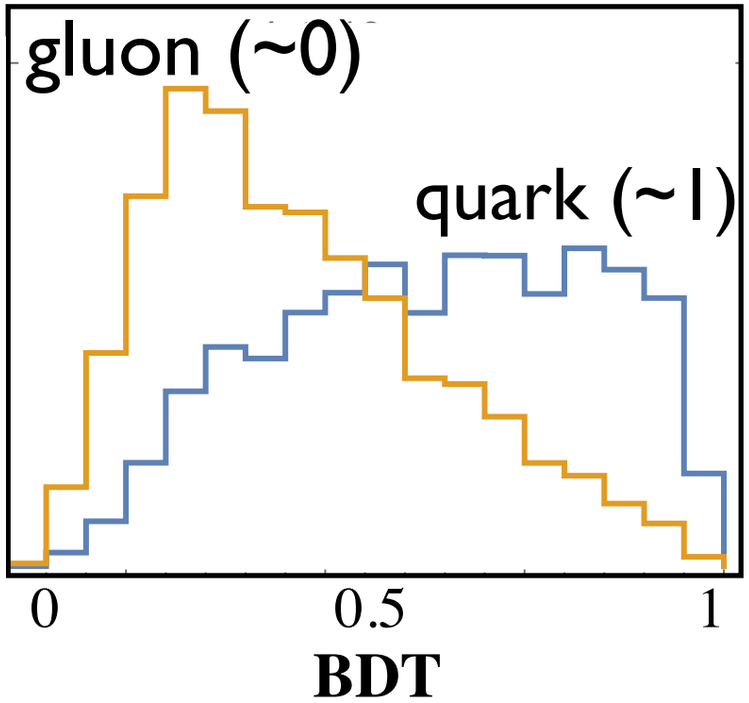
all jets are quark jets

$$\frac{S}{B} \propto \frac{\epsilon_S}{\epsilon_B} \sim \left(\frac{\epsilon_q}{\epsilon_g}\right)^{N_q^{signal} - N_q^{QCD-bkg}}$$





MC analysis



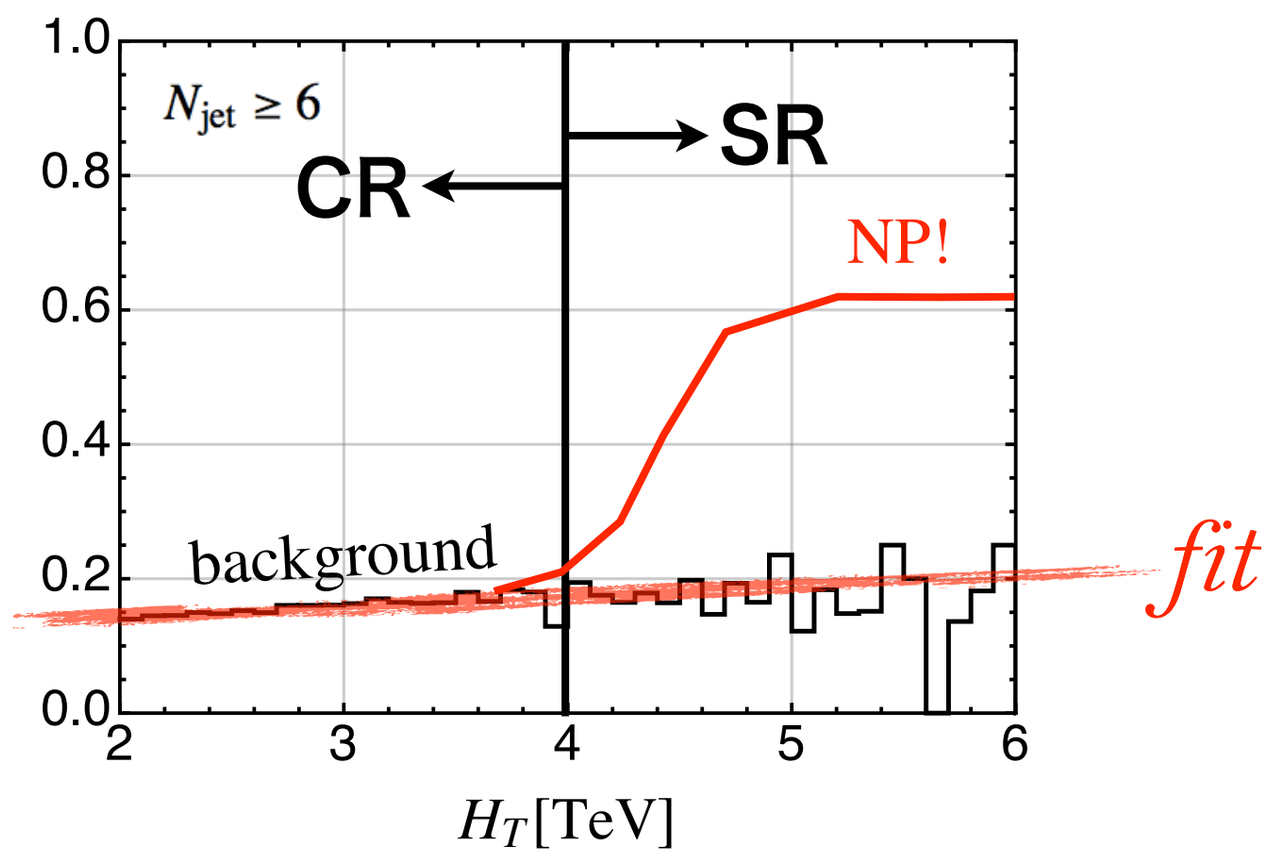
QCD jets $\overrightarrow{\mathrm{BDT}} \sim (1, 0, 0, 0, 0, \ldots), \ (1, 1, 0, 0, 0, \ldots)$

signal $\overrightarrow{\mathrm{BDT}} \sim (1, 1, 1, 1, 1, \ldots)$

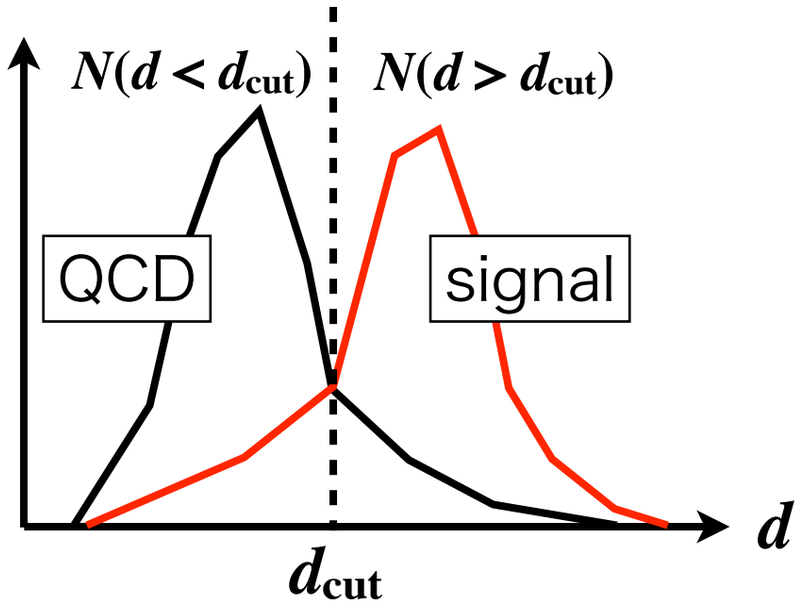
- BDT distance: $d = \sqrt{\left|\overrightarrow{\mathrm{BDT}}\right|^2}$

- We can estimate # of background of each bins by data-driven extrapolations

$$\frac{N(d > d_{\mathrm{cut}})}{N}$$



MC analysis

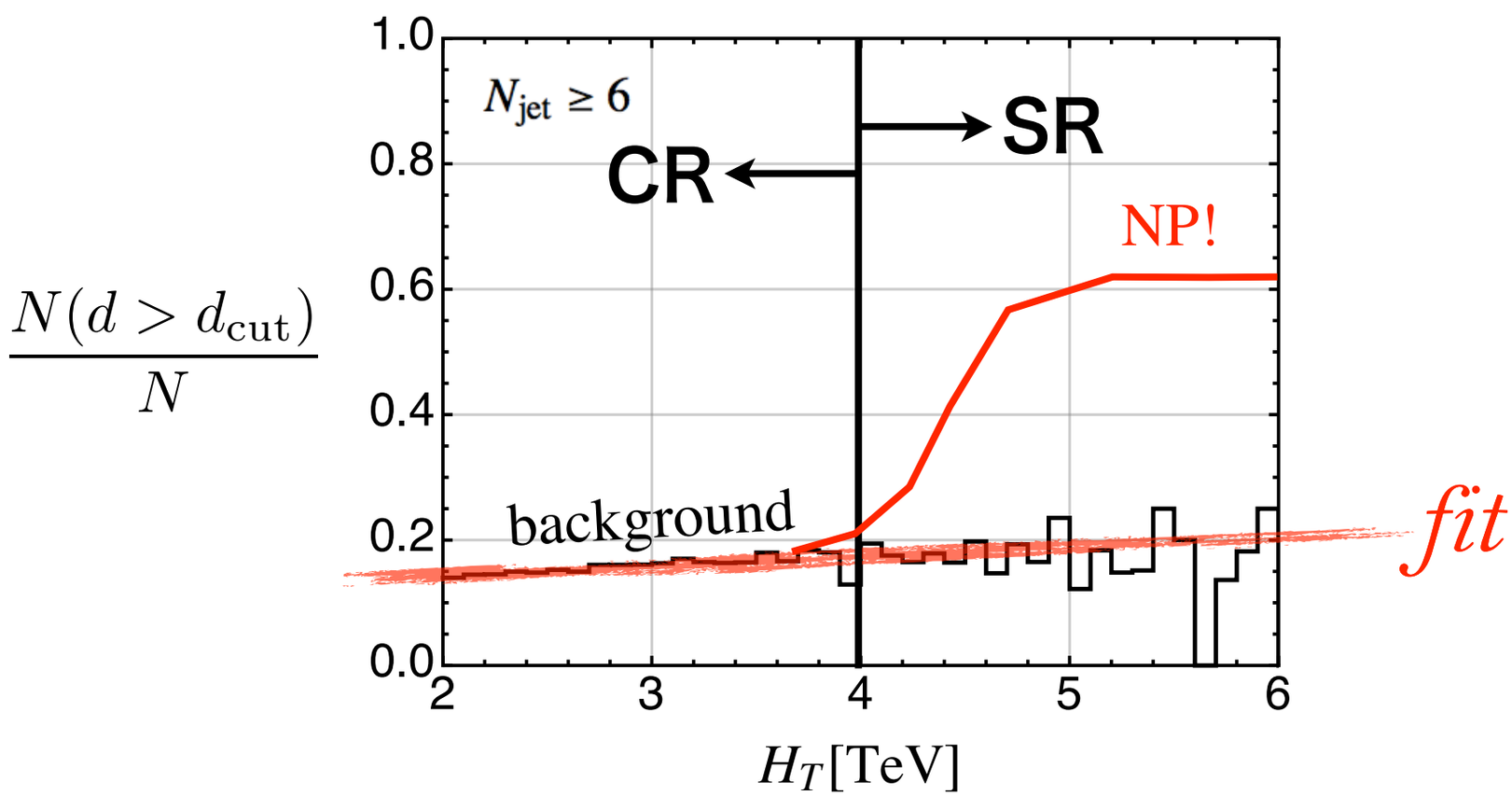


QCD jets $\overrightarrow{\mathrm{BDT}} \sim (1, 0, 0, 0, 0, \ldots), \ (1, 1, 0, 0, 0, \ldots)$

signal $\overrightarrow{\mathrm{BDT}} \sim (1, 1, 1, 1, 1, \ldots)$

- BDT distance: $d = \sqrt{\left|\overrightarrow{\mathrm{BDT}}\right|^2}$

- We can estimate # of background of each bins by data-driven extrapolations



Enhancement of S/B with d

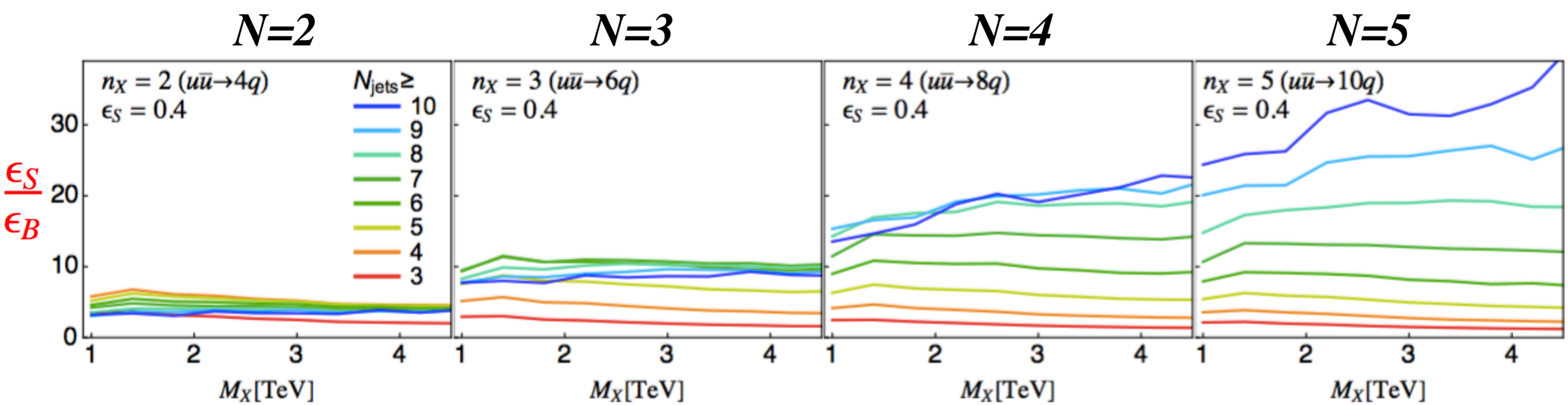
selection cut — jet substructure cut (d cut)

$$\frac{S}{B} \sim \frac{\sigma_S}{\sigma_B} \times \left(\frac{\epsilon_S}{\epsilon_B}\right) \times \left(\frac{\epsilon_S}{\epsilon_B}\right) \times \left(\frac{\epsilon_S}{\epsilon_B}\right)$$

H_T cut

toy-signal: $gg/q\bar{q} \to XX,\ X \to N$-quarks

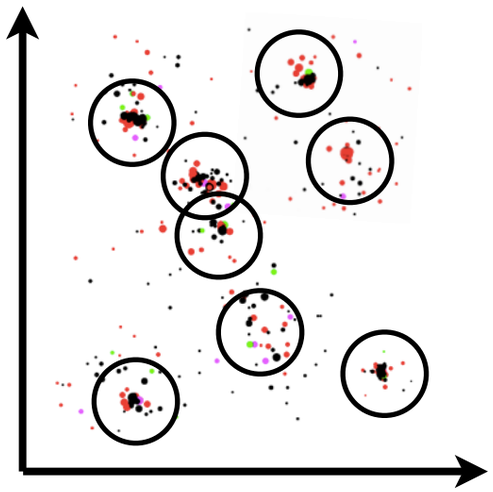
After imposing H_T cut. Fixed at $\epsilon_S = 0.4$



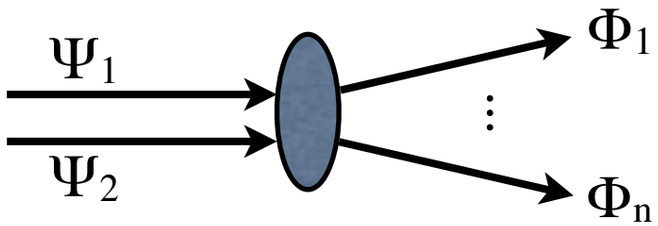
- Large enhancement of S/B

Summary

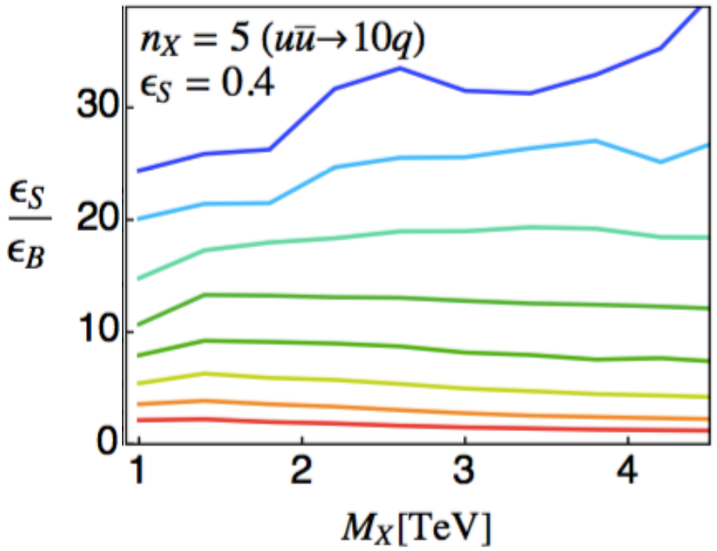
- Quark/gluon discrimination can be maximally utilized for BSM searches in multi-jet final states.



- Quark and gluon jet fraction in QCD multi-jet background was estimated.



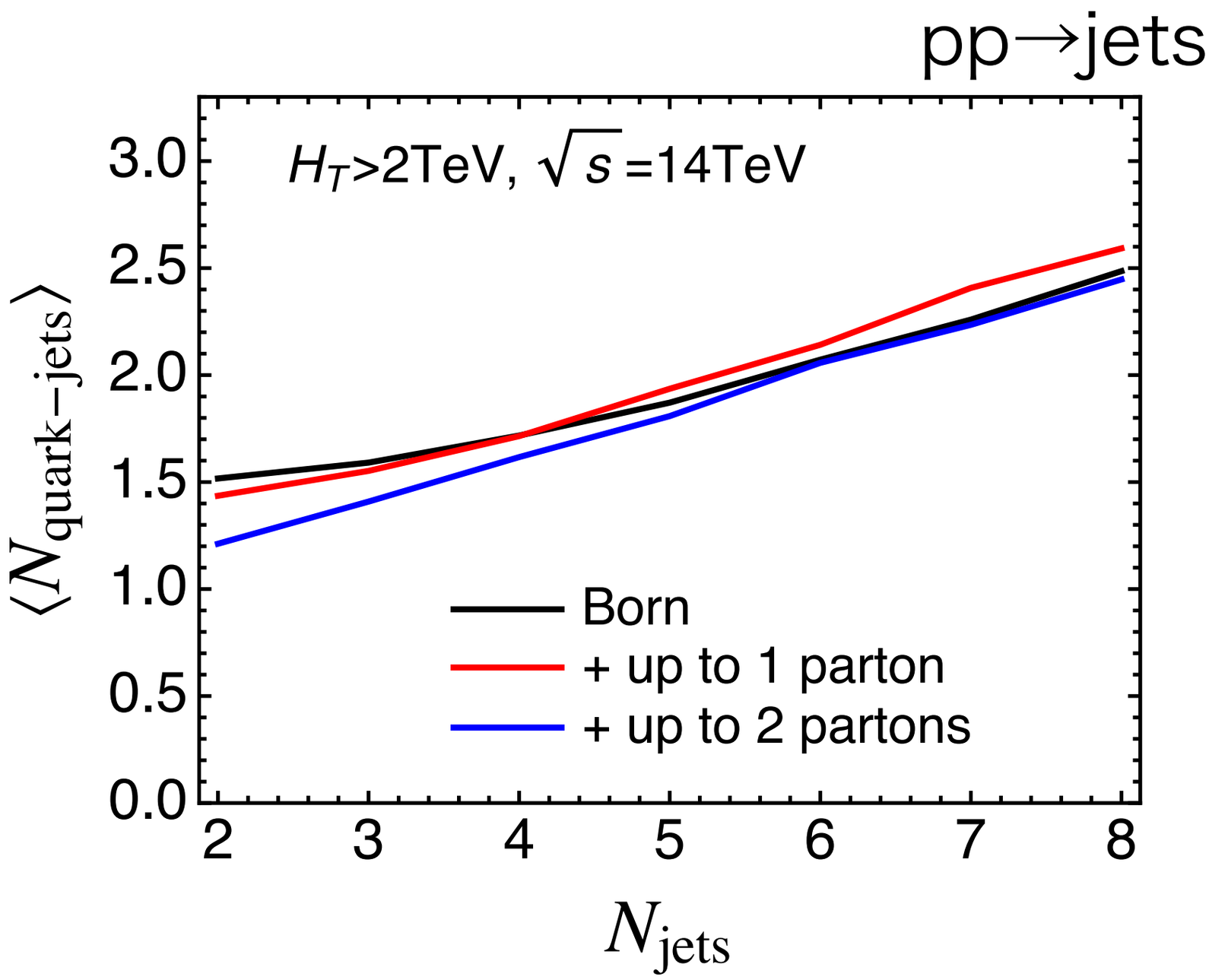
- Introducing a variable for the data-driven analysis in multi-jet final states, we checked the large improvement of S/B using the variable.



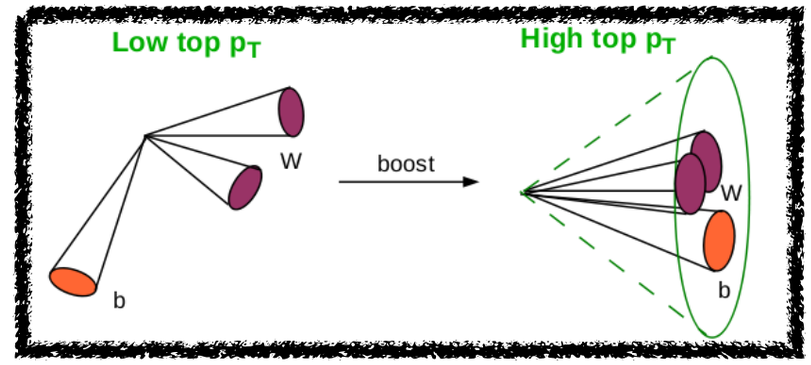
Back up

Belle II

Matrix element correction



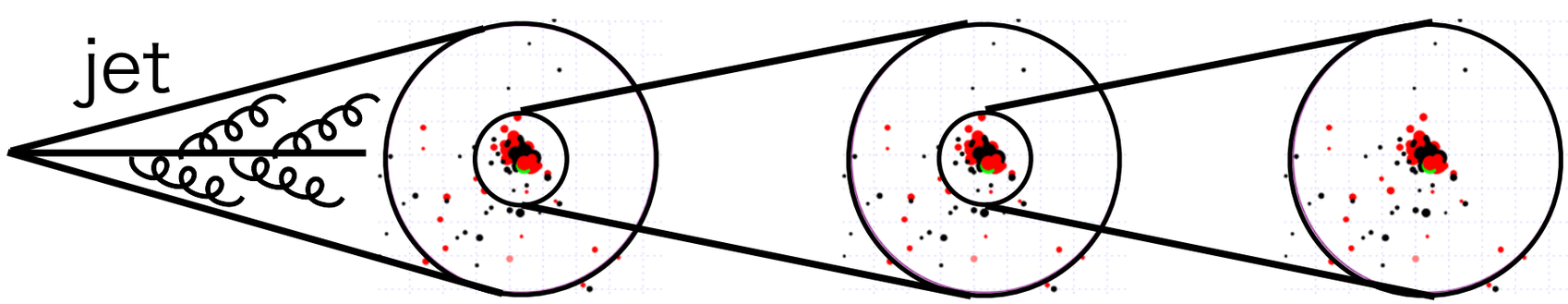
- Jet substructure technique was established well as top/W/Z/H tagging tools (2-, 3-prong structure)



- Quark/Gluon discrimination is also available (1-prong structure)

jet substructure	formed by	R (jet radius)
top W Z H	EW	~1.0 (fat jet)
quark gluon	QCD	0.4

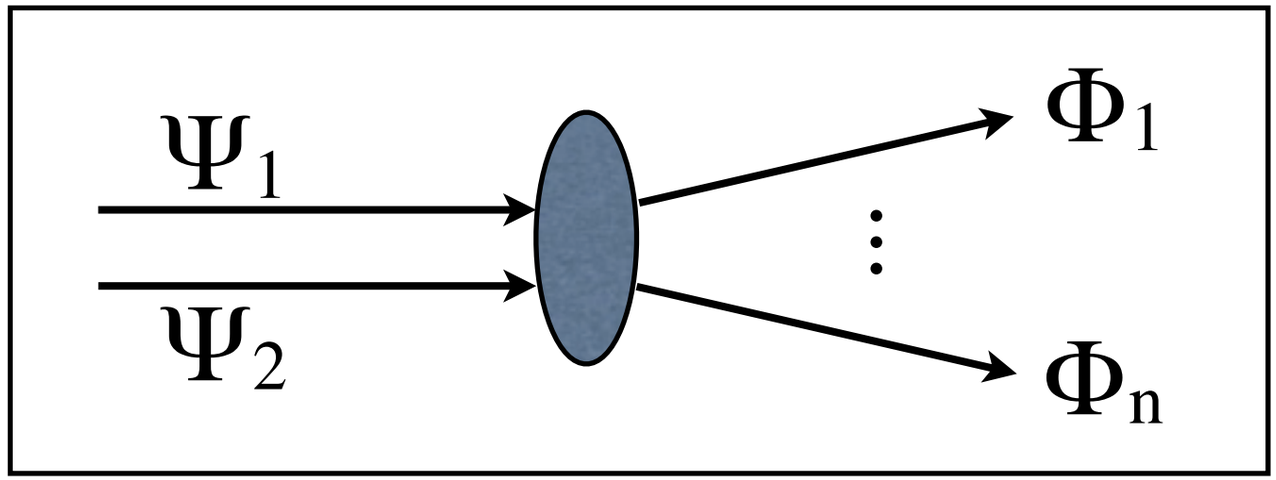
- QCD radiation is approximately scale invariant



- Quark/Gluon discrimination works well even small-R (even for R<0.4)

ISR+FSR

$W =$



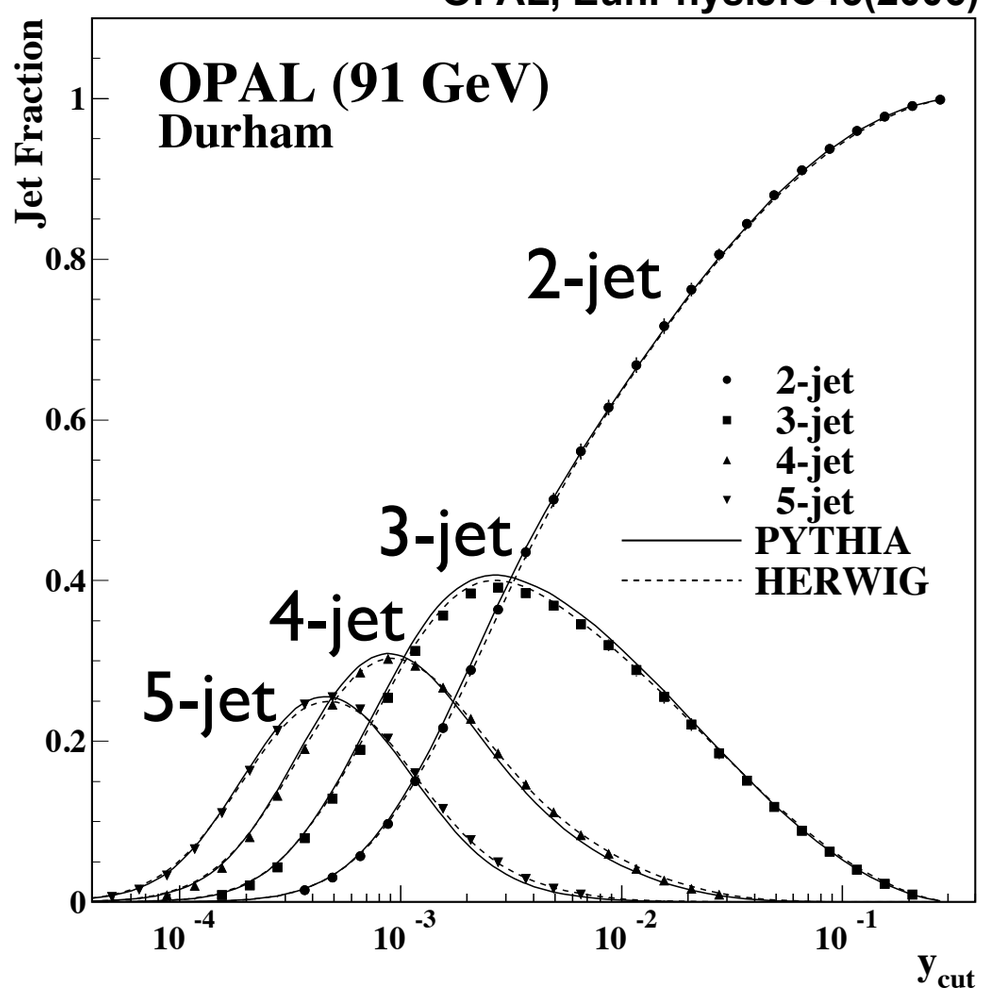
$$W(t) = \Psi_1 \Psi_2 \left(\prod_k \Phi_k \right) \qquad R_{n,m}(t) = \frac{1}{n!\, m!} \frac{\partial^n}{\partial u^n} \frac{\partial^m}{\partial v^m} W(t) \,|_{u=v=0}$$

- A whole generating functional for a matrix element is given by a product of FSR and ISR generating functionals.

Jet rates

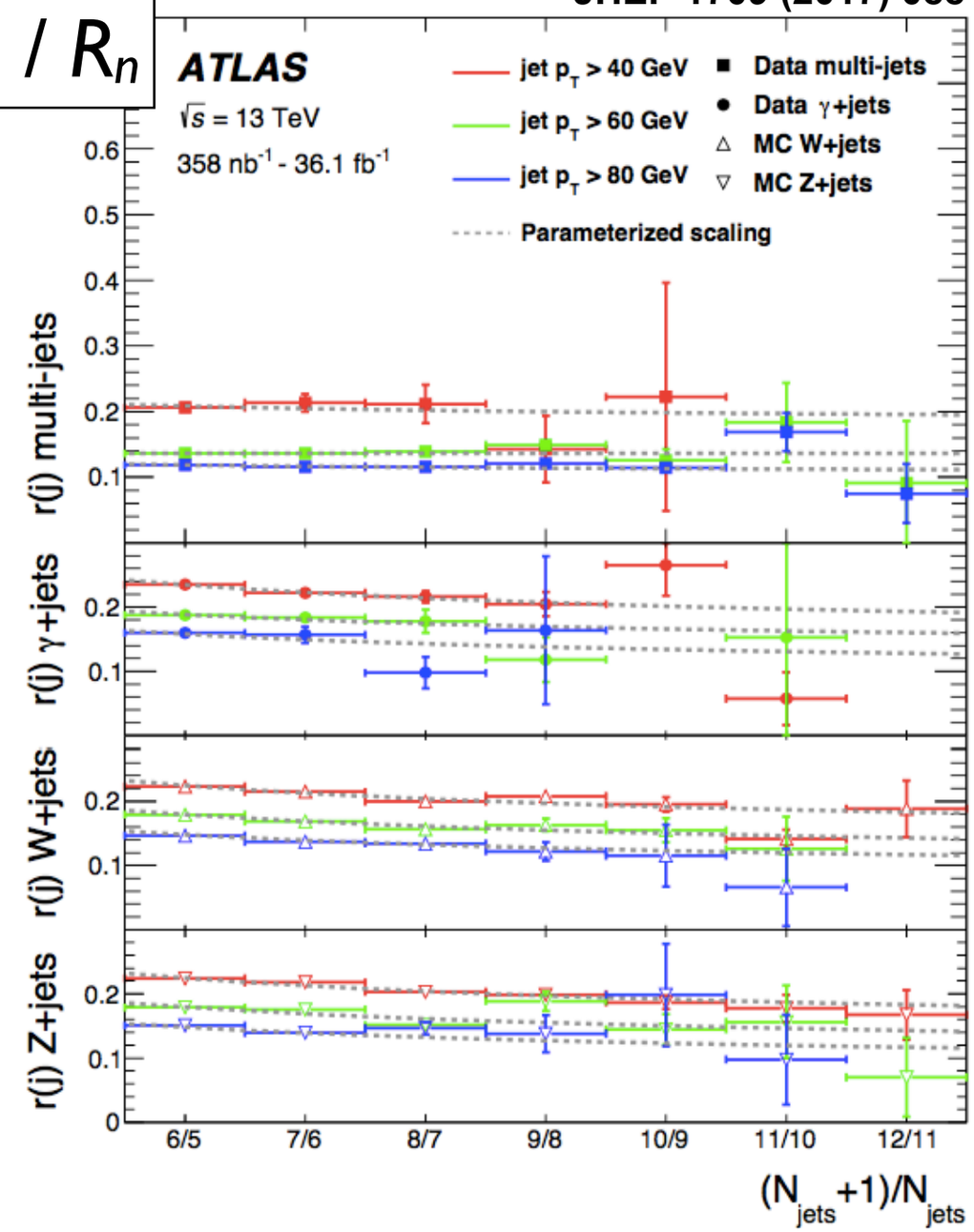
- $R_n(t)$: Probability that an event has *n* jet
- Studied well. Contribute to understanding of QCD
- and used ...

OPAL, Eur.Phys.J.C45(2006)

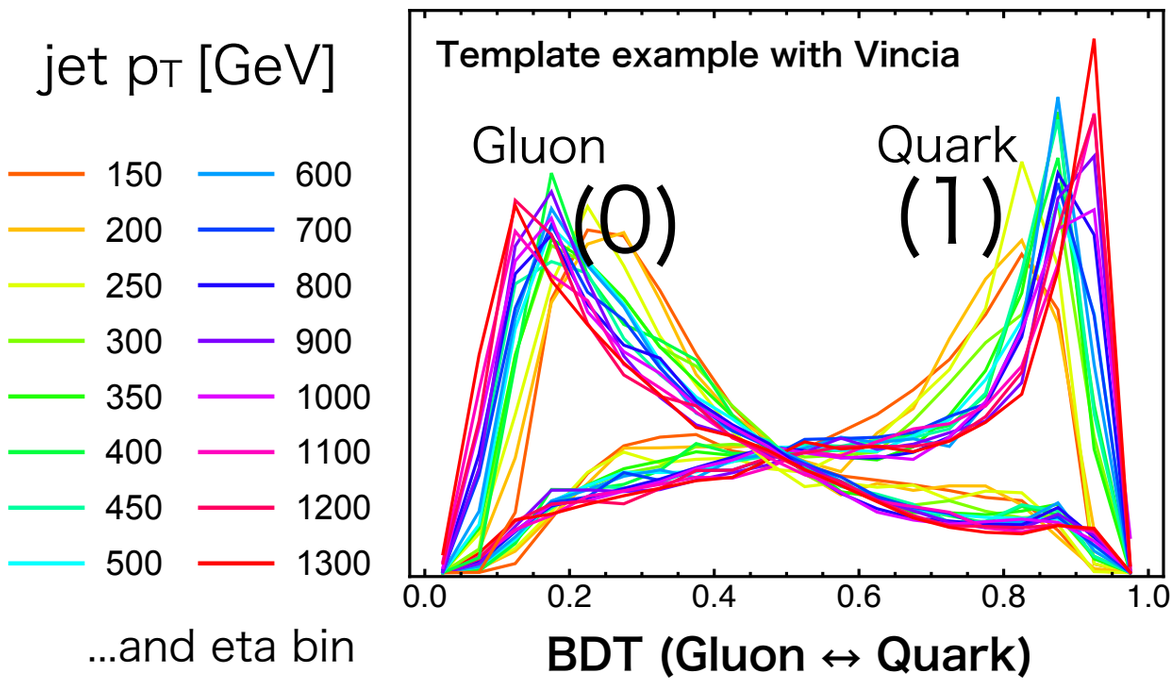


R_{n+1} / R_n

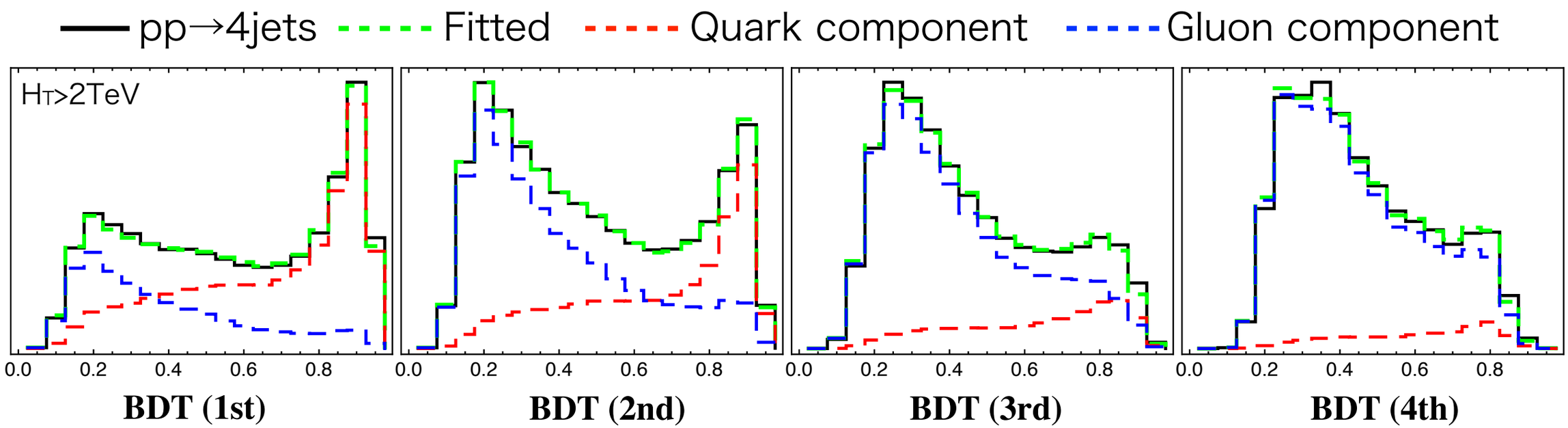
JHEP 1709 (2017) 088



How to measure quark jet rates in multi-jet final state

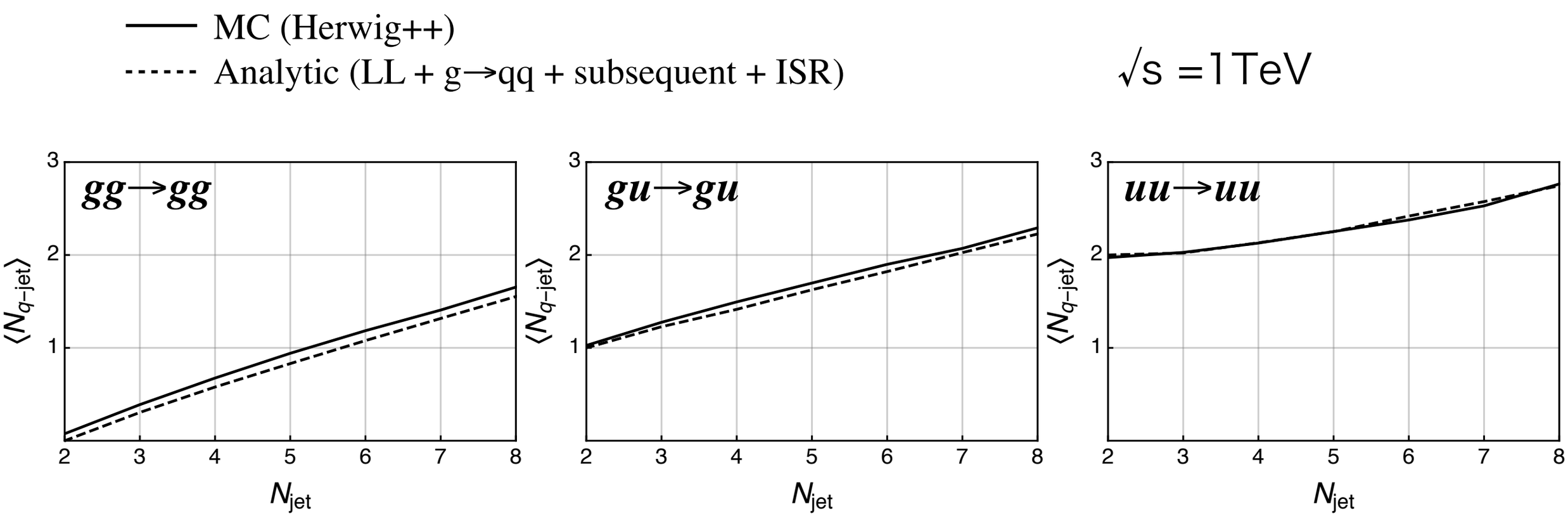


- Make quark/gluon jet templates for a variable from di-jet, Z/γ+jet sample



- Measurable, if the QCD jet substructure is universal (It depends on only p_T and rapidity, not # of jet)
- Many applications are conceivable

MC comparison



- Increment of jet (leading, LL), quark jet (sub-leading, NLL)
- QCD jets background is composed of 1 or 2 valence quark jets and many gluon jets