Jet Substructure at the LHC

Wouter Waalewijn







DESY - May 19, 2014

Outline

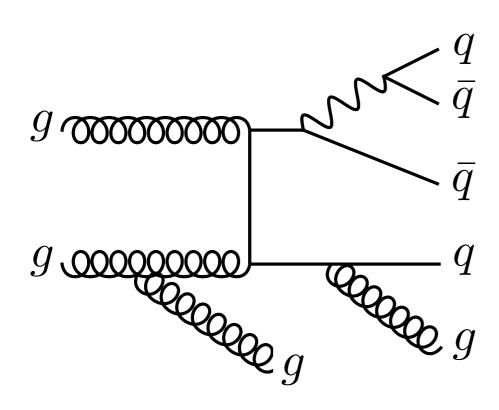
- Introduction
- Jet Charge
- Track-Based Observables
- Jet Mass
- Hadronization of Jets
- Conclusions

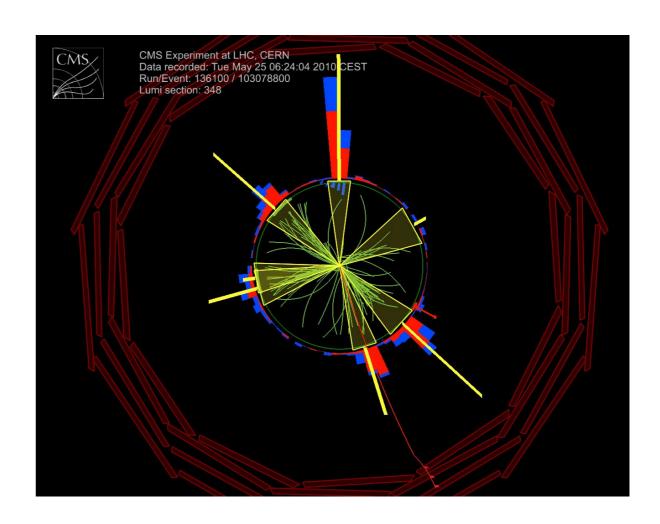


Introduction

What is a Jet?

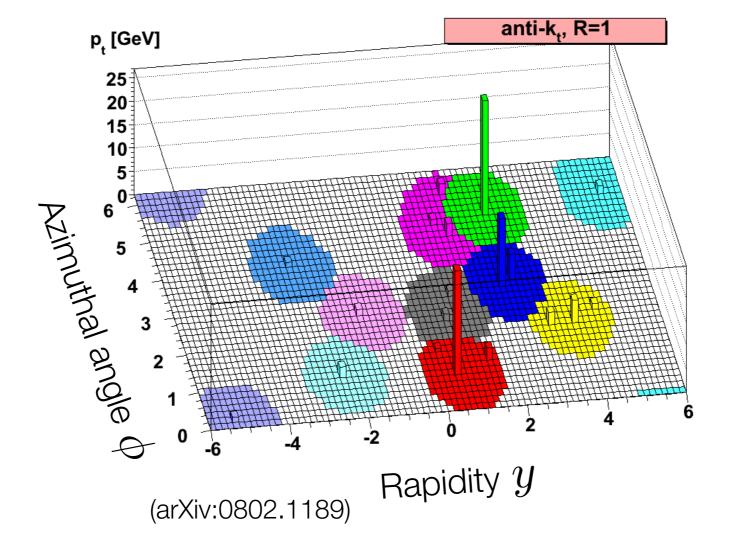
→ Produce jets of hadrons





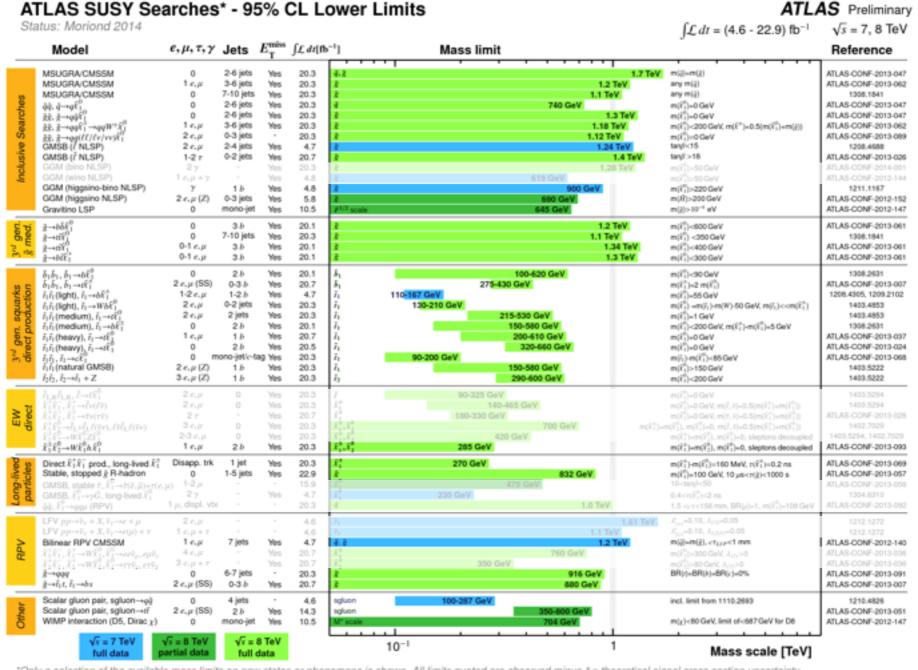
Jet Algorithms

- Repeatedly cluster nearest "particles" $p_i, p_j \rightarrow p_i + p_j$
- Cut off by jet radius R
- Default at LHC: anti- k_T (Cacciari, Salam, Soyez)



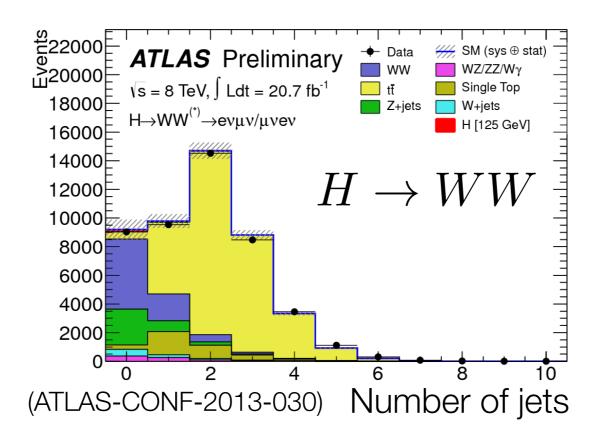
Jets at the LHC

Most measurements involve jets as signal or background



Jet Cross Sections

Bin by jet multiplicity to improve background rejection



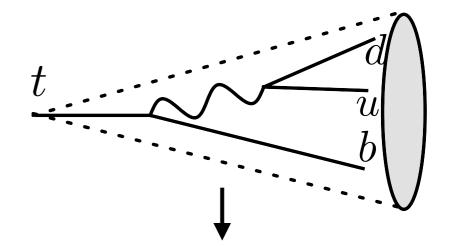
Large logarithms lead to large theory uncertainties

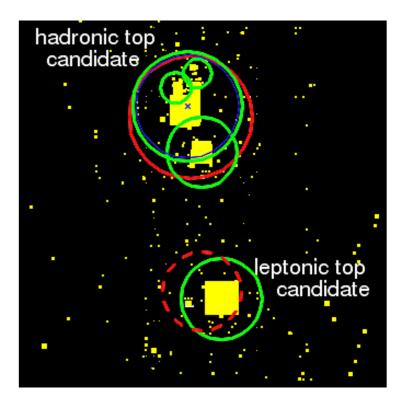
$$\sigma(H+0 \text{ jets}) \propto 1 - \frac{6\alpha_s}{\pi} \ln^2 \frac{p_T^{\text{cut}}}{m_H} + \dots$$

(Berger, Marcantonini, Stewart, Tackmann, WW; Banfi, Monni, Salam, Zanderighi, Becher, Neubert, Rothen; Stewart, Tackmann, Walsh, Zuberi; Liu, Petriello; Boughezal, Liu, Petriello, Tackmann, Walsh; Bernlocher, Gangal, Gillberg, Tackmann, ...)

Jet Substructure for Boosted Objects

• Decay products of boosted t, W, H can lie within one jet

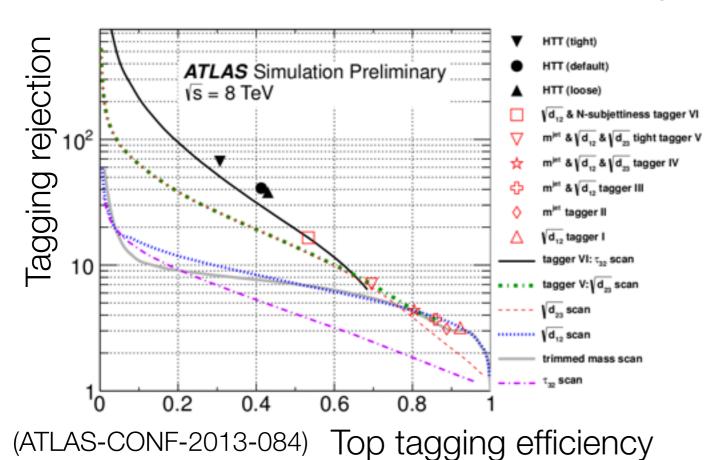


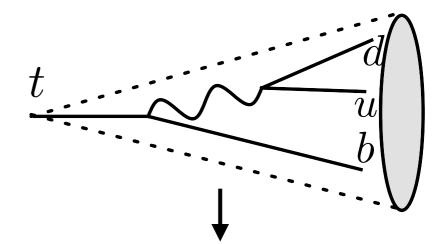


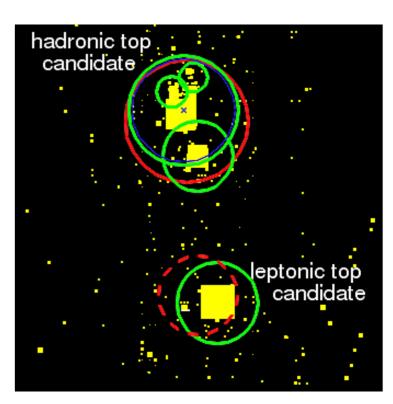
(ATLAS-CONF-2013-052)

Jet Substructure for Boosted Objects

- Decay products of boosted t, W, H can lie within one jet
- It started with mass drop for $H \to bb$ (Butterworth, Davison, Rubin, Salam)
- Plethora of substructure techniques







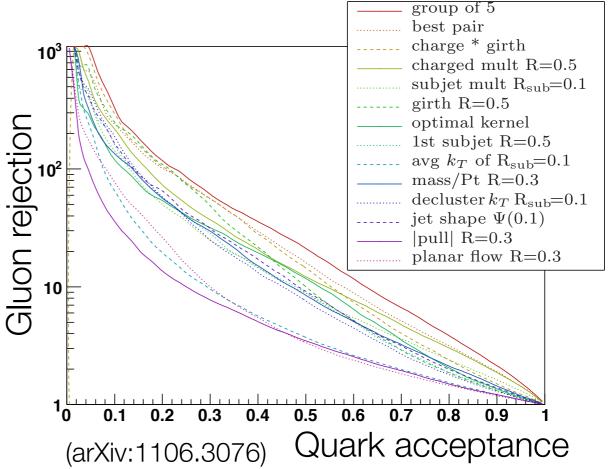
(ATLAS-CONF-2013-052)

Jet Substructure for Quark/Gluon Discrimination

- New physics often quark, QCD backgrounds often gluon
- Extensive Pythia study (Gallicchio, Schwartz)
 - Charged track multiplicity and jet "girth" are good

girth =
$$\sum_{i \in \text{jet}} \frac{p_T^i}{p_T^J} \sqrt{(y_i - y_J)^2 + (\phi_i - \phi_J)^2}$$

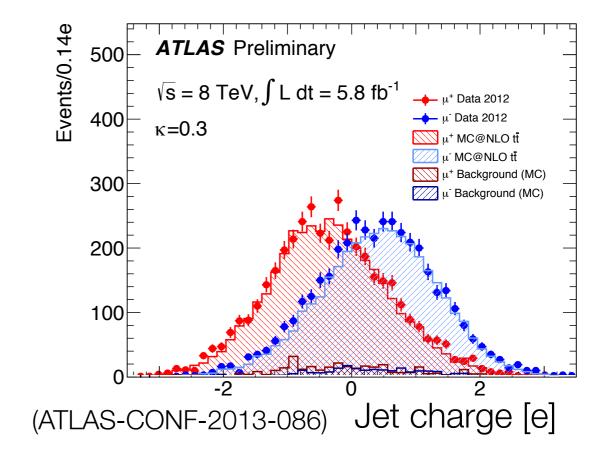
 More variables only give marginal improvement

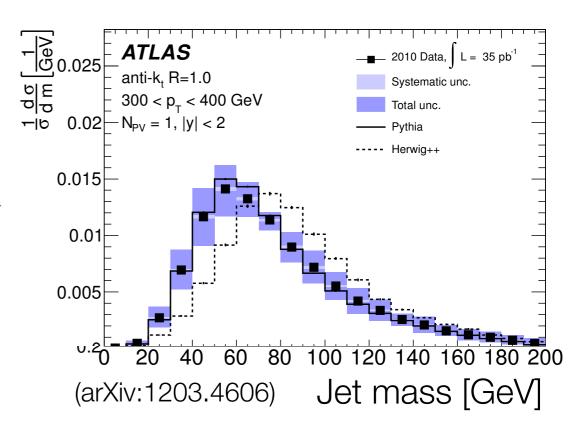


Jet Mass and Charge

Motivation:

- Measured at the LHC
- Benchmark for our ability to calculate substructure
- Test and improve Monte Carlo: Herwig and Pythia differ





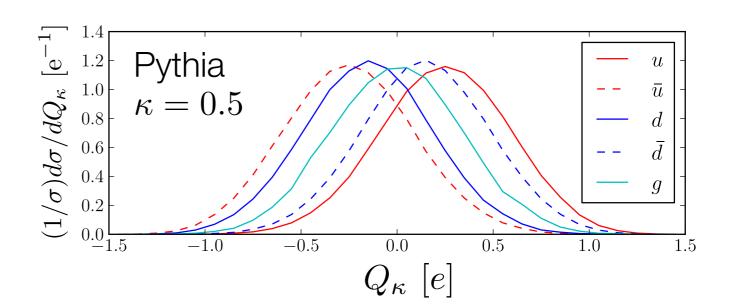
Jet Charge

Krohn, Lin, Schwartz, WW (arXiv:1209.2421)

WW (arXiv:1209.3091)

Defining Jet Charge

$$Q_{\kappa} = \sum_{i \in \text{jet}} Q_i \Big(\frac{p_T^i}{p_T^J}\Big)^{\kappa}$$
 (Feynman, Field)



If κ too small:

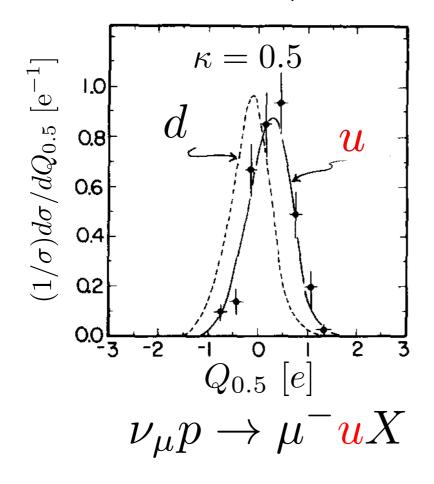
- Sensitive to soft hadrons → contamination
- $\kappa=0$ similar to multiplicity (Recent work by Bolzoni, Kniehl, Kotikov)

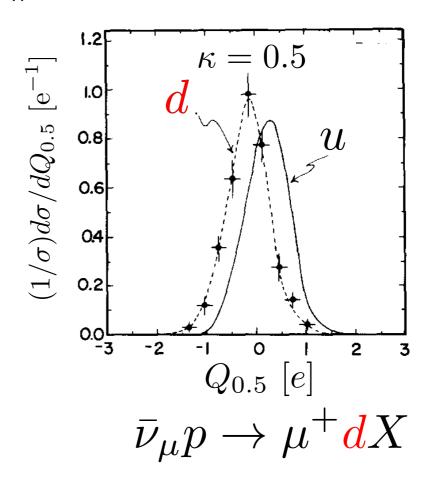
If κ too large:

Only sensitive to leading hadron → need more statistics

Historical Applications

Test parton model (Fermilab (1980))



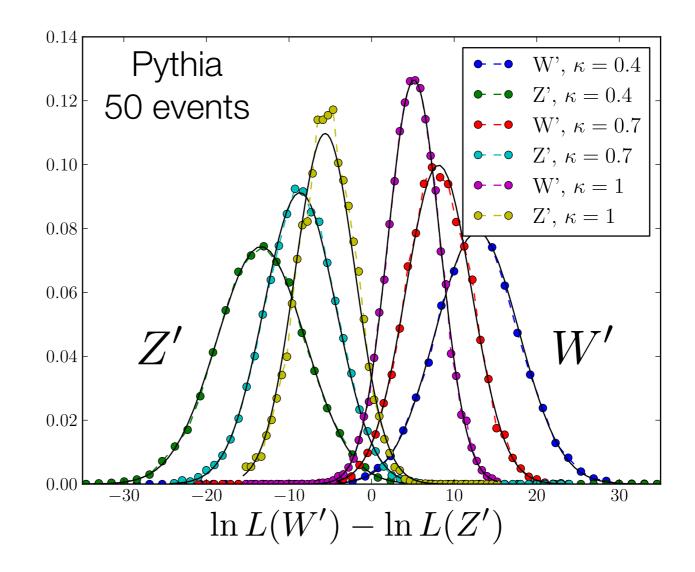


- Jet charge at LEP:
 - Forward-backward charge asymmetry (AMY (1990),...)
 - $B^0 \leftrightarrow \overline{B^0}$ mixing (ALEPH (1992), ...)

Possible LHC application: W' vs. Z'

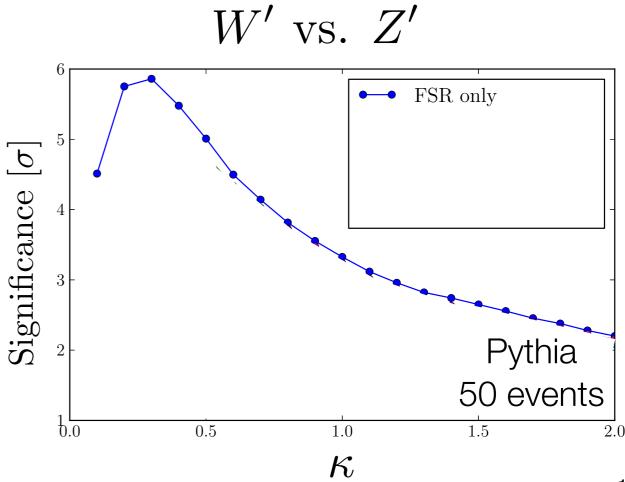
- Hadronically decaying W' or Z' with 1 TeV mass
- 2-dim. likelihood discriminant based on both jet charges

$$Z' o u ar u$$
 $Z' o d ar d$
 $Z' o d ar d$
 $VS.$
 $W' o u ar d$
 $W' o d ar u$



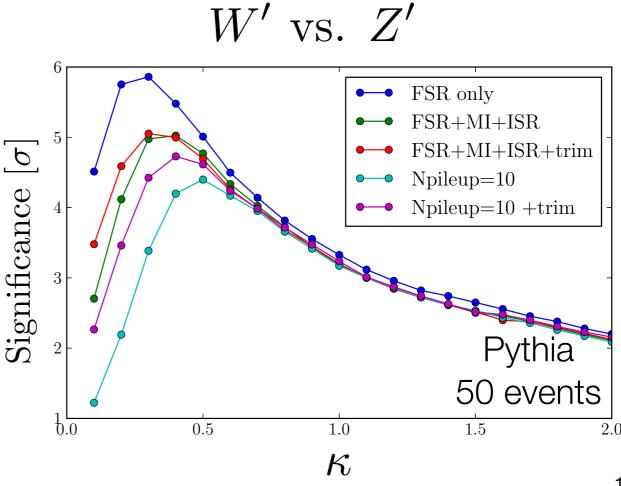
LHC Challenges

- Trade off between soft contamination and statistics
- · We did not include: backgrounds, detector effects, ...



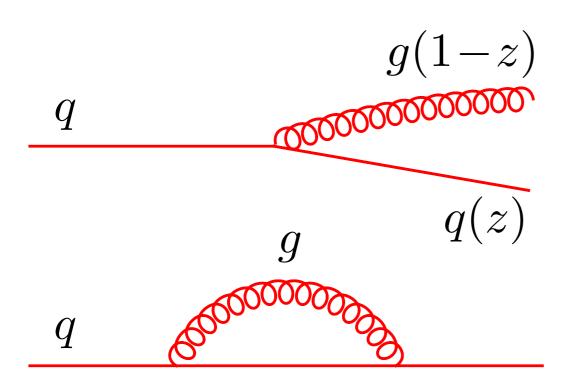
LHC Challenges

- Trade off between soft contamination and statistics
- We did not include: backgrounds, detector effects, ...
- Various sources of contamination:
 - Initial State Radiation
 - Multiparton Interactions
 - Pile-up (overestimated)
- All soft \rightarrow increase κ



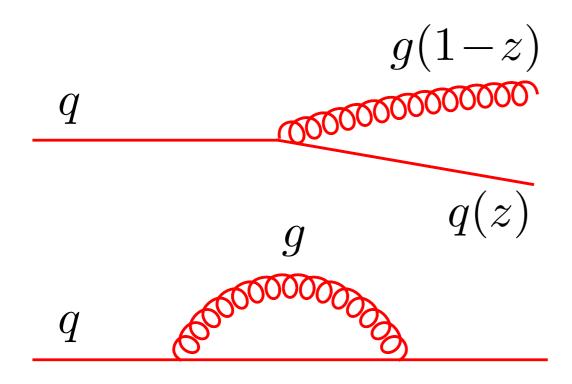
Jet Charge Not IR Safe

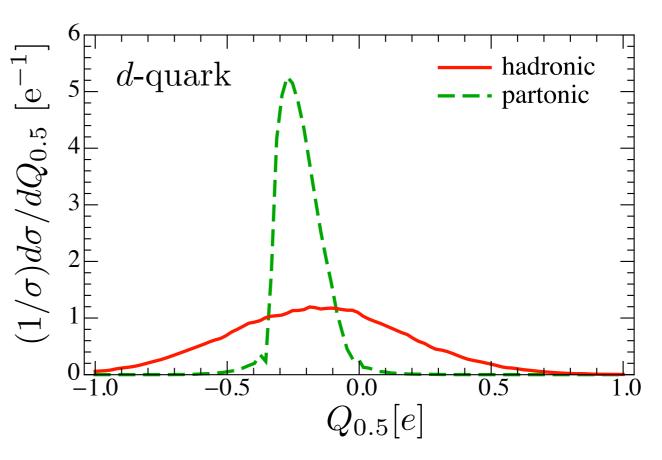
- Consider $q \rightarrow qg$ in collinear limit
- $Q_q z^{\kappa} \neq Q_q \rightarrow$ divergences don't cancel between real/virtual



Jet Charge Not IR Safe

- Consider $q \to qg$ in collinear limit
- $Q_q z^{\kappa} \neq Q_q \rightarrow$ divergences don't cancel between real/virtual
- Jet charge only defined for hadrons





Average Jet Charge Calculation

$$\langle Q_{\kappa} \rangle = \sum_{h} \int dz \ Q_{h} z^{\kappa} \ \frac{1}{\sigma_{\text{jet}}} \frac{d\sigma_{h \in \text{jet}}}{dz}$$
hadron h charge weight

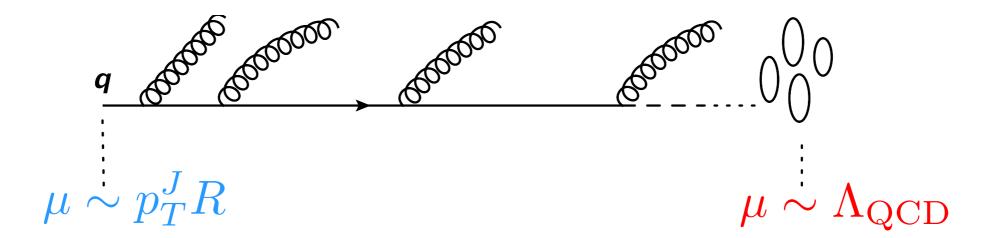
• At LO, weight = fragmentation function $D_q^h(z,\mu \sim p_T^J R)$

Jet scale

Average Jet Charge Calculation

$$\langle Q_{\kappa} \rangle = \underbrace{\sum_{h} \int dz}_{\text{hadron } h} \underbrace{Q_{h} z^{\kappa}}_{\text{charge}} \underbrace{\frac{1}{\sigma_{\text{jet}}} \frac{d\sigma_{h \in \text{jet}}}{dz}}_{\text{weight}}$$

- At LO, weight = fragmentation function $D_q^h(z,\mu \sim p_T^J R)$
- Calculate p_T^J, R dependence from evolution to $\mu \sim \Lambda_{
 m QCD}$
- $D_q^h(z, \mu \sim \Lambda_{\rm QCD})$ describes hadronization

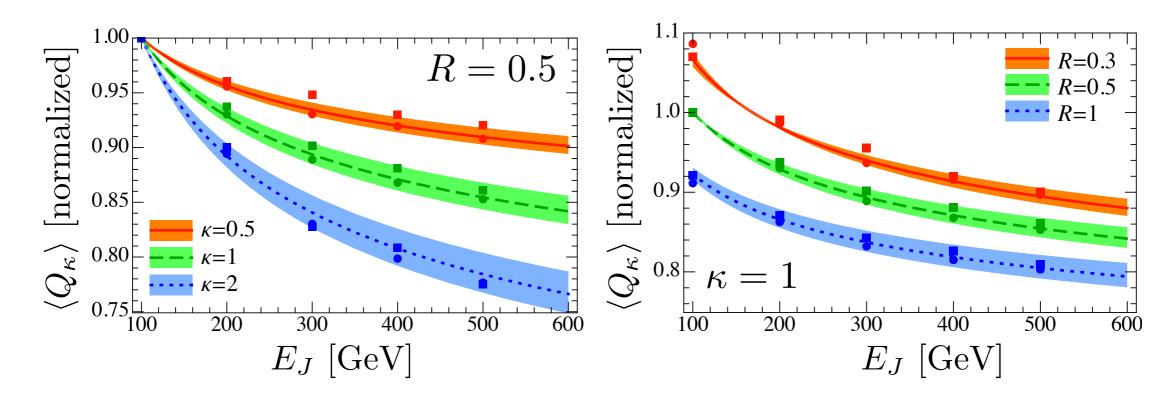


Jet scale

DGLAP Evolution vs. Pythia's Shower

$$\langle Q_{\kappa}(E_J R, \text{flavor}) \rangle = \text{perturbative}(\kappa, E_J R) \times \text{hadronization}(\kappa, \text{flavor})$$
perturbative splitting + evolution

- Normalize average jet charge: $\frac{\langle Q_{\kappa}(E_JR)\rangle}{\langle Q_{\kappa}(50~{
 m GeV})\rangle}$
 - → Hadronization (and flavor dependence) drops out



Good agreement (including perturbative splitting)

Fragmentation Functions vs. Pythia's Hadronization

• Average jet charge at $E_J = 100 \text{ GeV}, R = 0.5$

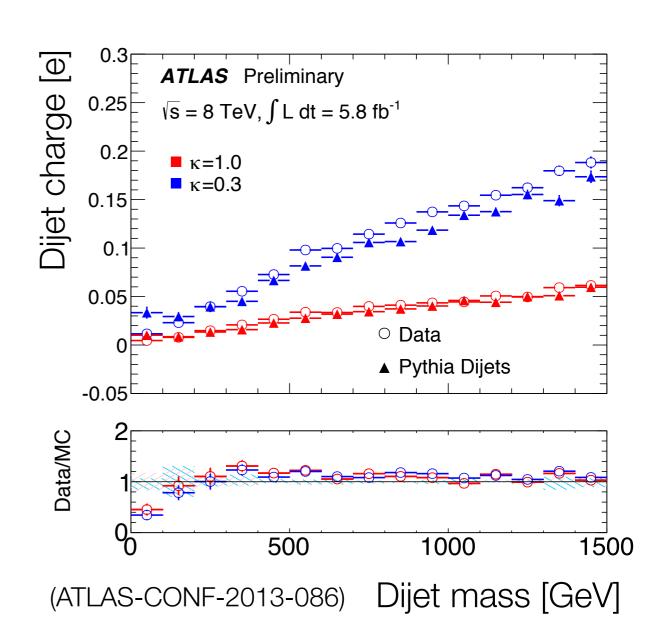
	$oldsymbol{u}$ -quark			<i>d</i> -quark		
κ	PYTHIA	DSS	AKK08	Pythia	DSS	AKK08
0.5	0.271	0.237	0.221	-0.162	-0.184	-0.062
1	0.144	0.122	0.134	-0.078	-0.088	-0.046
2	0.055	0.046	0.064	-0.027	-0.030	-0.027

(DSS = De Florian, Sassot, Stratmann, AKK08 = Albino, Kniehl, Kramer)

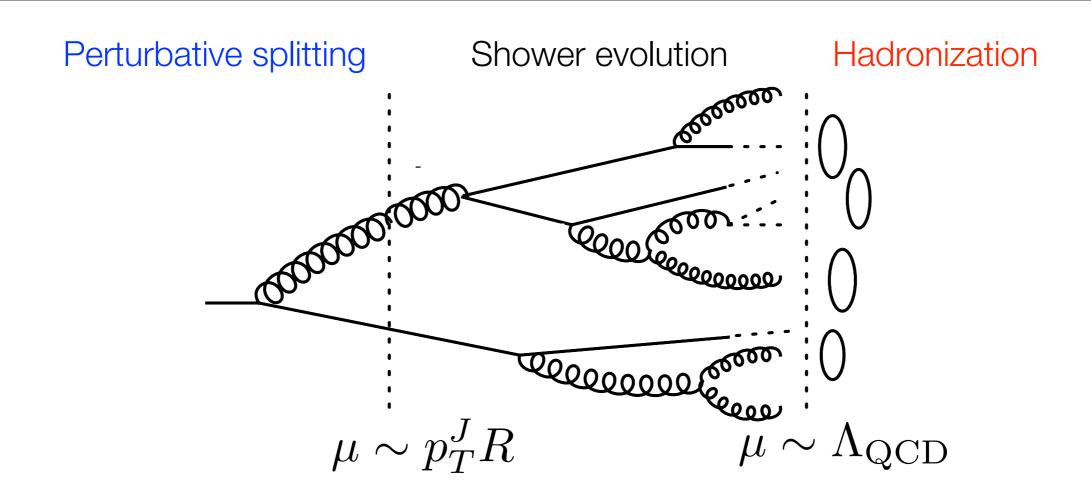
- Pythia consistent with fragmentation functions
- Large uncertainties as we need $D_q^{h^+}-D_q^{h^-}=D_q^{h^+}-D_{\bar q}^{h^+}$ Most fragmentation data is e^+e^- giving $D_q^{h^+}+D_{\bar q}^{h^+}$

Average Dijet Charge at the LHC

- Charge increases with dijet mass due to PDFs
- Pure QCD measurement of valence structure of proton!

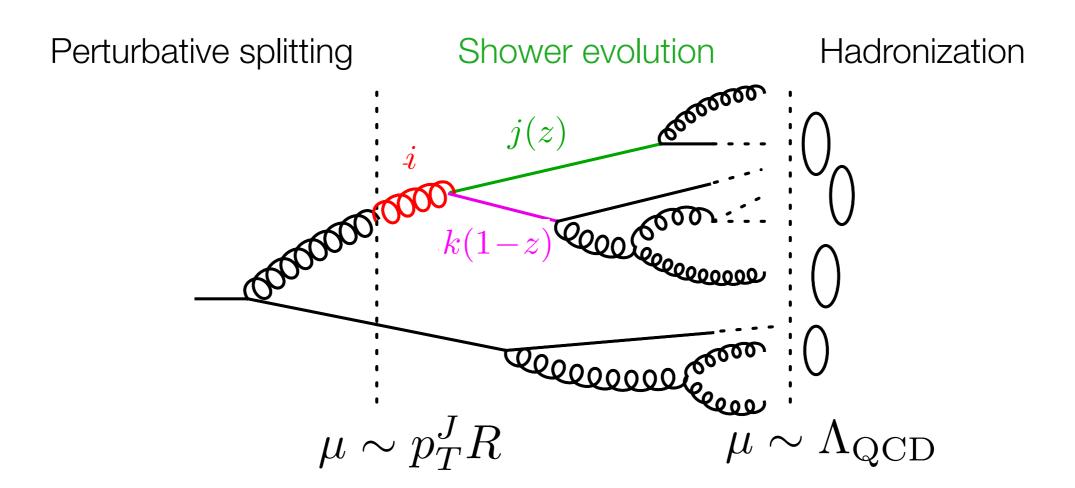


Full Jet Charge Distribution



- Perturbative splitting reduces μ -dependence (Jain, Procura, WW)
- Hadronization depends on full charge distribution $D_i(Q_\kappa,\mu)$
 - Similar to multi-hadron fragmentation function

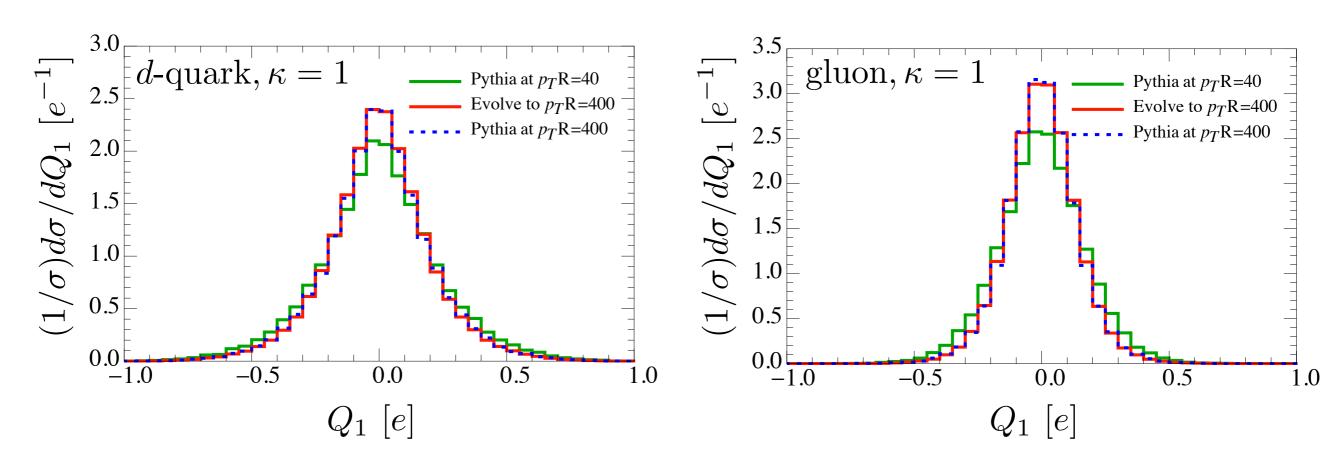
Full Jet Charge Distribution



Splitting probability Sample over distributions of branches
$$\mu \frac{d}{d\mu} D_i(Q_{\kappa}, \mu) = \sum_j \int dz \, \frac{\alpha_s}{2\pi} P_{ji}(z) \int dQ_{\kappa}^a D_j(Q_{\kappa}^a, \mu) \int dQ_{\kappa}^b D_k(Q_{\kappa}^b, \mu) \times \delta[Q_{\kappa} - z^{\kappa}Q_{\kappa}^a - (1-z)^{\kappa}Q_{\kappa}^b]$$
Charge is (weighted) sum of branches

DGLAP Evolution vs. Pythia's Shower

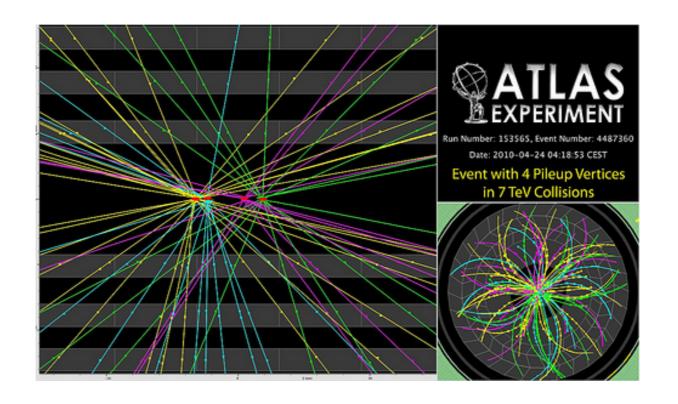
- Use Pythia as input and evolve, good agreement
- Distribution changes much more slowly than PDF or FF



Track-Based Observables

Chang, Procura, Thaler, WW (arXiv:1303.6637, 1306.6630)

Track-Based Observables



- Advantages:
 - better angular resolution
 less sensitive to pile-up
- Disadvantages:
 - smearing of resonance peaks
 not collinear safe

Formalism

Calculating an IR safe observable e:

$$\frac{d\sigma}{de} = \sum_{N} \int d\Pi_{N} \, \frac{d\sigma_{N}}{d\Pi_{N}} \, \delta[e - \hat{e}(\{p_{i}^{\mu}\})]$$

• Corresponding track-based observable \bar{e} :

$$\frac{d\sigma}{d\overline{e}} = \sum_{N} \int d\Pi_{N} \underbrace{\frac{d\overline{\sigma}_{N}}{d\Pi_{N}}} \underbrace{\int \prod_{i=1}^{N} dx_{i} \, T_{i}(x_{i})}_{\text{hadronization}} \delta[\overline{e} - \hat{e}(\{x_{i}p_{i}^{\mu}\})]$$

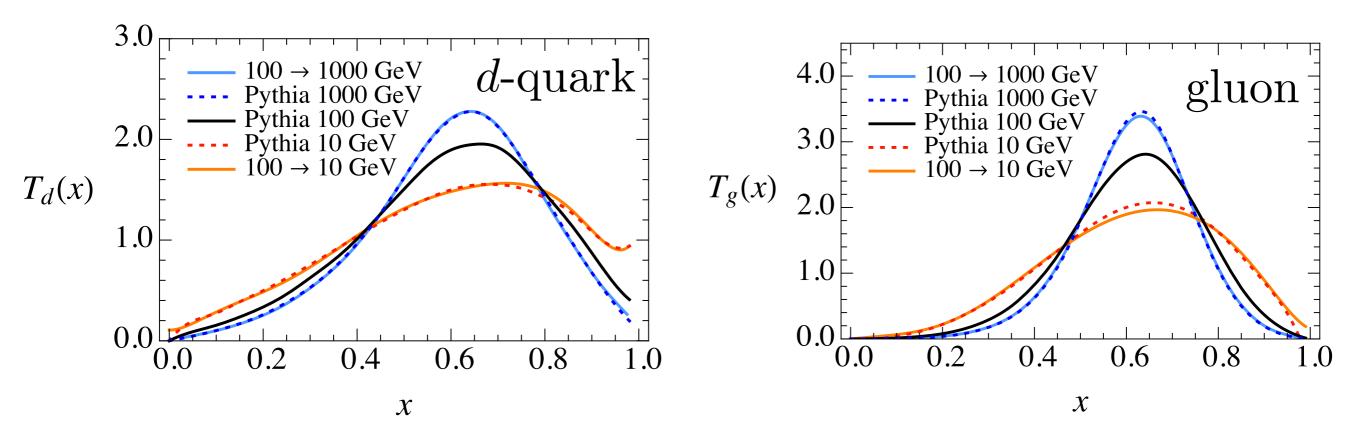
$$\text{IR div. removed} \qquad \text{conversion to tracks:}$$

$$\overline{p}_{i}^{\mu} = x_{i}p_{i}^{\mu} + \mathcal{O}(\Lambda_{\text{QCD}})$$

 Similar to matching PDFs but now one track function for each parton

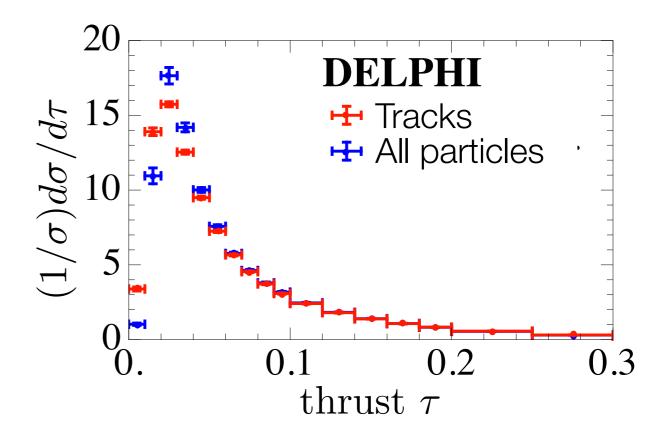
Track Functions $T_i(x, \mu)$

- Describes average energy fraction x converted to tracks
- Large width from hadronization fluctuations (smearing)
- Similar μ -evolution as jet charge, agrees with Pythia



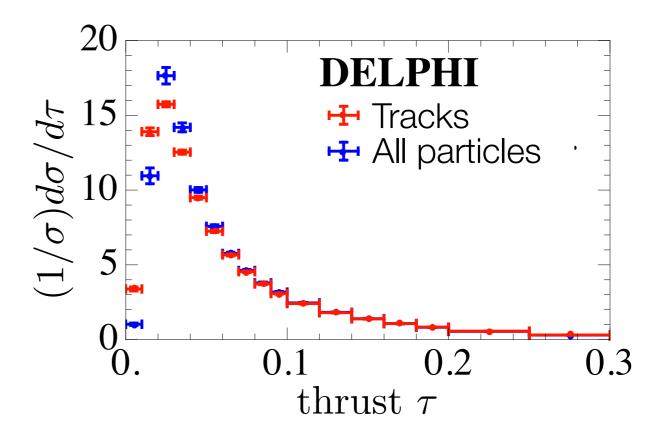
Track-based thrust

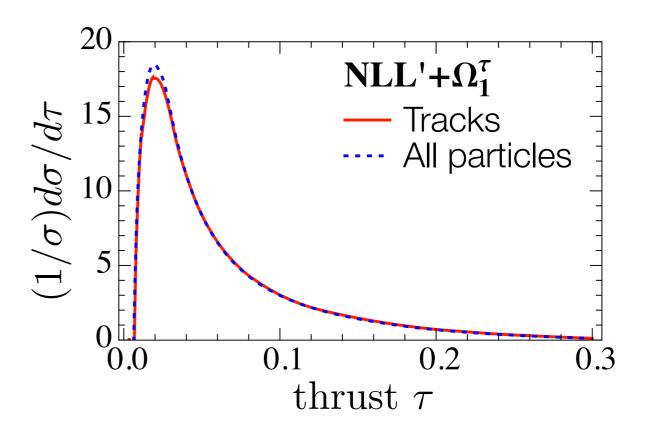
 Measured at LEP using charged and all particles



Track-based thrust

- Measured at LEP using charged and all particles
- We find same except in non-perturbative peak
- Smearing of peak is small for dimensionless ratios





Jet Mass

Jouttenus, Tackmann, Stewart, WW (arXiv:1302.0846) Tackmann, Stewart, WW (to appear)

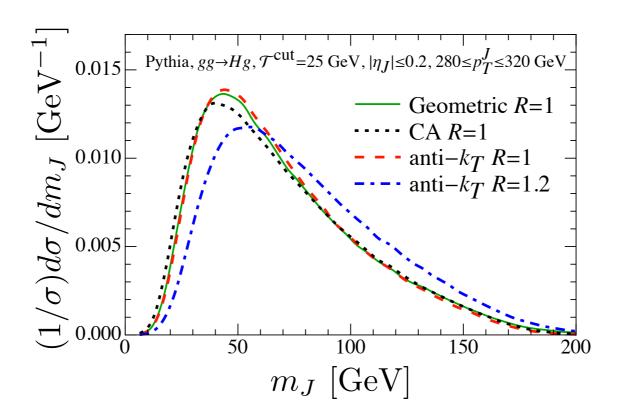
Jet Mass Resummation

- Jet mass is defined as $\, m_J^2 = \left(\sum_{i \in \mathrm{jet}} p_i^\mu \right)^2 \,$
- Cross section contains logarithms of $L = \ln(m_J^{\rm cut}/p_T^J)$

$$\int_{0}^{m_{J}^{\text{cut}}} dm_{J} \frac{d\sigma}{dm_{J}} = \sigma_{0} \left\{ 1 + \alpha_{s} \left[c_{12}L^{2} + c_{11}L + c_{10} + n_{1} \left(m_{J}^{\text{cut}} \right) \right] \right. \\ \left. + \alpha_{s}^{2} \left[c_{24}L^{4} + c_{23}L^{3} + c_{22}L^{2} + c_{21}L + c_{20} + n_{2} \left(m_{J}^{\text{cut}} \right) \right] \right. \\ \left. + \alpha_{s}^{3} \left[c_{36}L^{6} + c_{35}L^{5} + c_{34}L^{4} + c_{33}L^{3} + c_{32}L^{2} + \dots \right] \right. \\ \left. + \left. \vdots \right. \right\} \\ \left. \text{LL} \quad \text{NLL} \quad \text{NNLL}$$

- Need to resum dominant higher-order effects for $m_J\!\ll\!p_T^J$
- Nonsingular n_i is suppressed by $(m_J^{\rm cut}/p_T^J)^2$

Jet Mass and Jet Definition



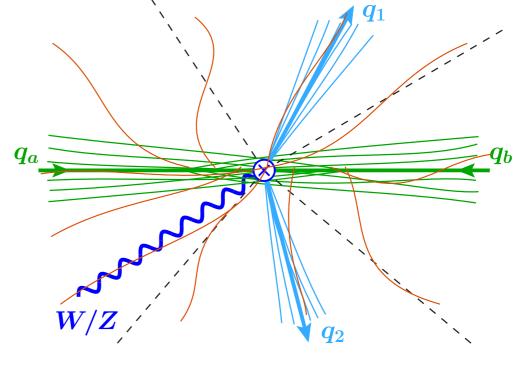
- Clustering algorithms theoretically complicated
- Jet mass spectrum is fairly independent of jet definition \rightarrow use N-jettiness (with correct R)

N-Jettiness Event Shape (Stewart, Tackmann, WW)

$$\mathcal{T}_N = \sum_i \min\{\hat{q}_a \cdot p_i, \hat{q}_b \cdot p_i, \hat{q}_1 \cdot p_i, \dots\}$$
 jet size parameter

- Reference vectors: $\hat{q}_{a,b}=(1,0,0,\pm 1)$, $\hat{q}_J=(1,\hat{n}_J)/\rho_J$
- $\mathcal{T}_N \to 0$ for exactly N jets, \mathcal{T}_N large for > N jets
- Very successfully used as substructure: N-Subjettiness

(Thaler, van Tilburg)

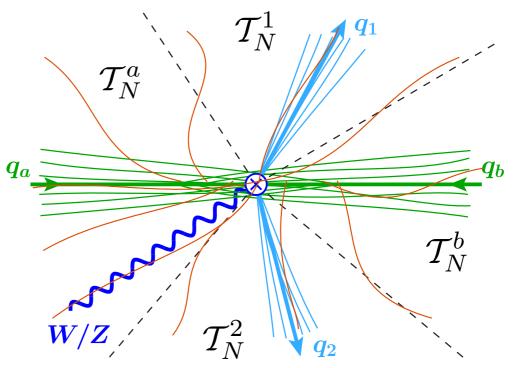


N-Jettiness Event Shape (Stewart, Tackmann, WW)

$$\mathcal{T}_N = \sum_{i} \min\{\hat{q}_a \cdot p_i, \hat{q}_b \cdot p_i, \hat{q}_1 \cdot p_i, \dots\} = \mathcal{T}_N^a + \mathcal{T}_N^b + \mathcal{T}_N^1 + \dots$$
beams jets

- Reference vectors: $\hat{q}_{a,b}=(1,0,0,\pm 1)$, $\hat{q}_J=(1,\hat{n}_J)/\rho_J$
- $\mathcal{T}_N \to 0$ for exactly N jets, \mathcal{T}_N large for > N jets
- Very successful substructure technique: N-Subjettiness (Thaler, van Tilburg)
- \mathcal{T}_N splits into contributions from each beam/jet region
- Related to jet mass:

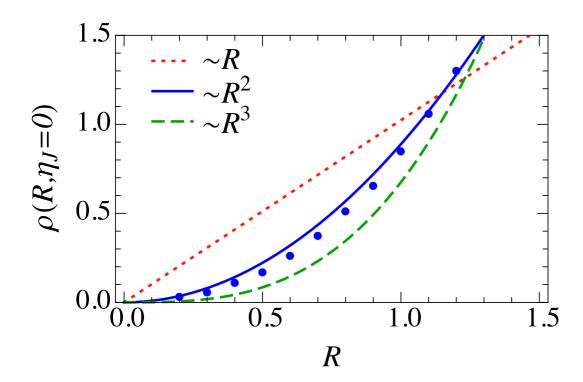
$$m_J^2 = 2\rho_J E_J \mathcal{T}_N^J$$

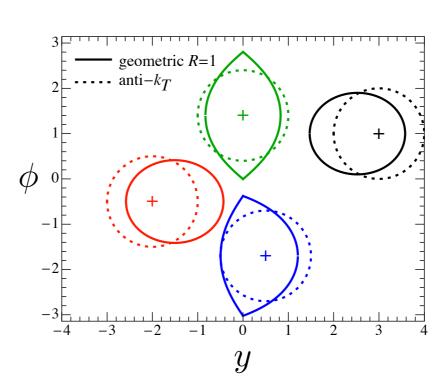


N-Jettiness Parameters

$$\mathcal{T}_N = \sum_{i} \min\{\hat{q}_a \cdot p_i, \hat{q}_b \cdot p_i, \hat{q}_1 \cdot p_i, \dots\} = \mathcal{T}_N^a + \mathcal{T}_N^b + \mathcal{T}_N^1 + \dots$$
beams jets

- Reference vectors: $\hat{q}_{a,b} = (1,0,0,\pm 1), \; \hat{q}_{J} = (1,\hat{n}_{J})/\rho_{J}$
- \hat{n}_J by minimizing \mathcal{T}_N or from jet alg. (same for $\mathcal{T}_N \to 0$)
- Choose $ho_J =
 ho(R, \eta_J)$ to match jet area of anti- k_T



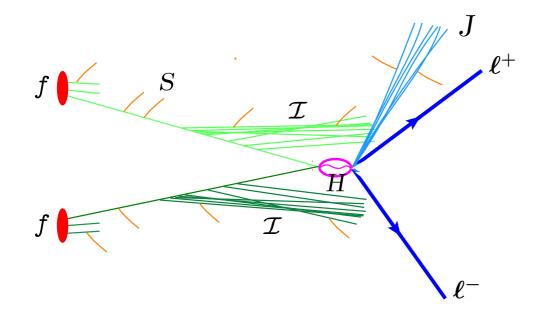


N-Jettiness Factorization

$$\frac{d\sigma(N \text{ jets})}{d\mathcal{T}_N^a d\mathcal{T}_N^b \cdots d\mathcal{T}_N^N} = \int dx_a \, dx_b \, d(\text{phase space}) \sum_{\kappa} \int dt_a \, B_{\kappa_a}(t_a, x_a, \mu)$$

$$\times \int dt_b \, B_{\kappa_b}(t_b, x_b, \mu) \prod_{J=1}^N \int ds_J \, J_{\kappa_J}(s_J, \mu) \, \operatorname{tr} \left[H_N^{\kappa}(\{q_i^{\mu}\}, \mu) \right]$$

$$\times S_N^{\kappa} \left(\mathcal{T}_N^a - \frac{t_a}{Q_a}, \mathcal{T}_N^b - \frac{t_b}{Q_b}, \dots, \mathcal{T}_N^N - \frac{s_N}{Q_N}, \{\hat{q}_i\}, \mu \right) \right]$$



- Hard scattering
- Initial state radiation (+PDFs)
- Final state radiation
- Soft radiation

N-Jettiness Factorization

- Separating physics at different scales enables resummation
- At NNLL order need one-loop B, J, H, S

B: Stewart, Tackmann, WW; Mantry, Petriello, J: Bauer, Manohar; Fleming, Leibovich, Mehen; Becher, Schwartz One-loop H for H+1-jet: Schmidt, One-loop S for N-jettiness: Jouttenus, Stewart, Tackmann, WW

Three-loop cusp and two-loop non-cusp anomalous dim.

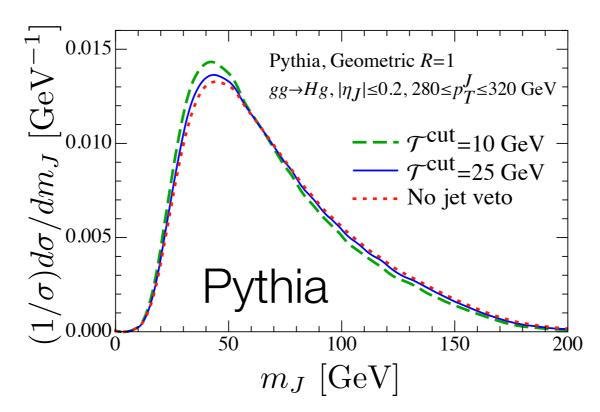
Three-loop cusp: Korchemsky, Radyushkin; Moch, Vermaseren, Vogt, Two-loop non-cusp known from: Kramer, Lampe; Harlander; Aybat, Dixon, Sterman; Becher, Neubert; Becher, Schwartz; Stewart, Tackmann, WW

Normalization

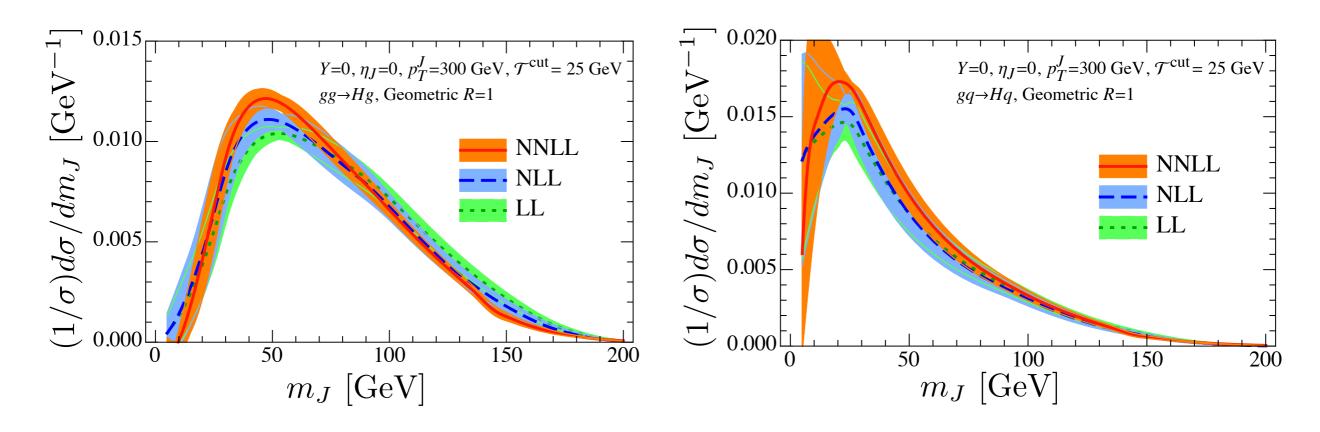
- We are required to veto additional jets through $\mathcal{T}_1^a, \mathcal{T}_1^b$
- Normalizing the spectrum removes this dependence:

$$\frac{\sigma(\mathcal{T}_1^a, \mathcal{T}_1^b \leq \mathcal{T}^{\text{out}}, m_J, p_T^J, y^J, Y)}{\int dm_J \, \sigma(\mathcal{T}_1^a, \mathcal{T}_1^b \leq \mathcal{T}^{\text{cut}}, m_J, p_T^J, y^J, Y)}$$

Experimental results are also normalized

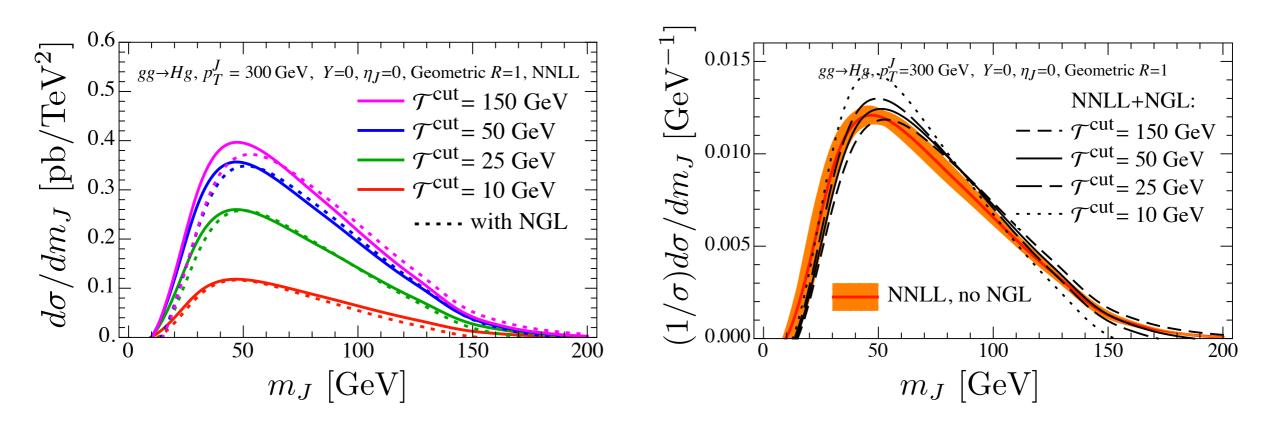


Perturbative Convergence



- We consider $gg \to Hg$ and $gq \to Hq$ (proxies for gluon and quark jets)
- Good agreement between LL, NLL, NNLL

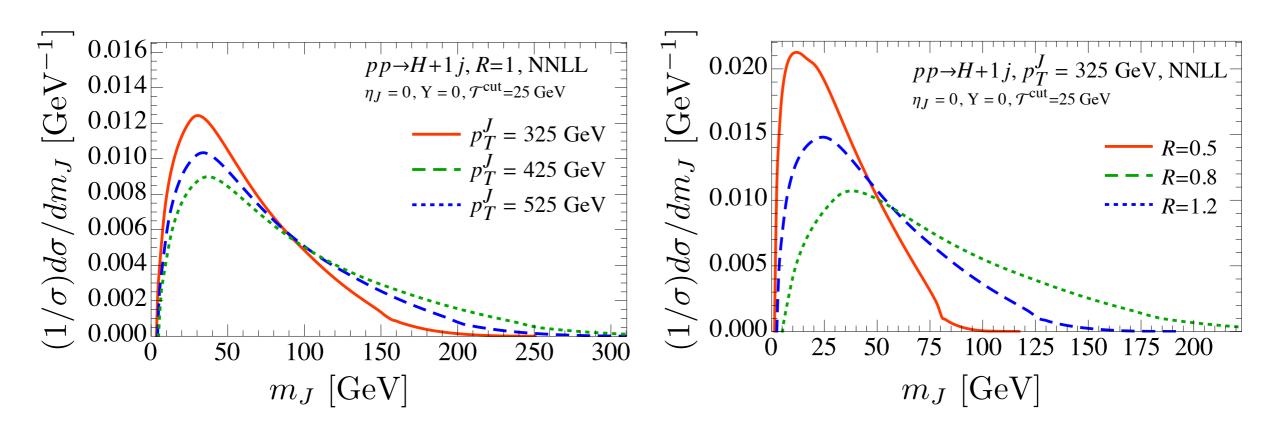
Non-Global Logarithms



- NGLs arise when soft probes multiple scales, $\mathcal{T}_N^i \ll \mathcal{T}_N^j$
- We find their effect to be small by testing with geometrical factor

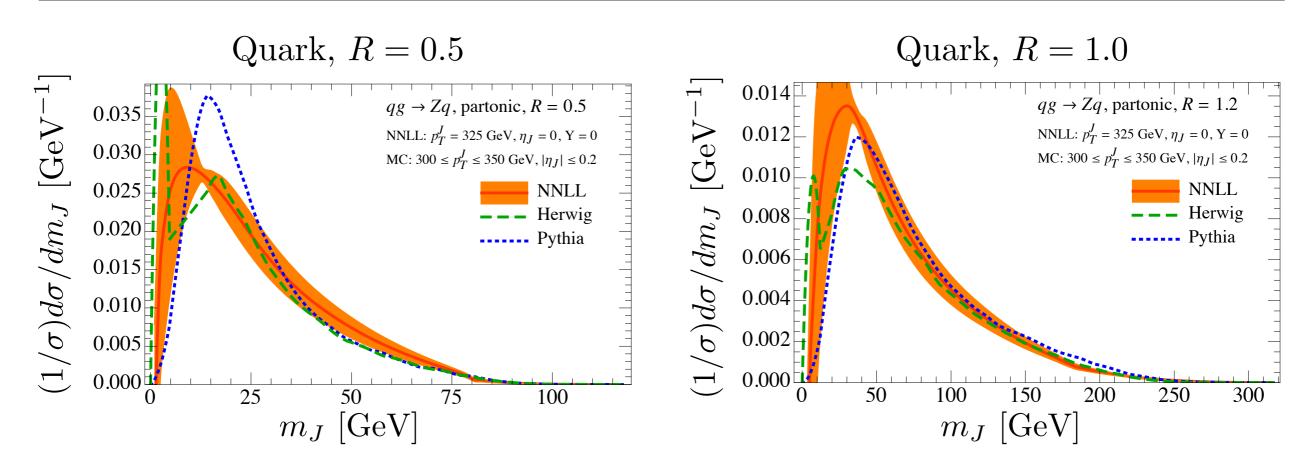
$$S_{\text{NGL}}(\{k_i^c\}, \mu_S) = \prod_i \left(\int_0^{k_i^c} dk_i \right) S_{\text{NGL}}(\{k_i\}, \mu_S) = -\frac{\alpha_s^2(\mu_S) C^2}{(2\pi)^2} \sum_{i < j} \sum_{i < j} G_{ij} \ln^2 \left(\frac{k_i^c}{k_j^c} \right)$$
color factor
$$C_{\text{NGL}}(\{k_i\}, \mu_S) = -\frac{\alpha_s^2(\mu_S) C^2}{(2\pi)^2} \sum_{i < j} \sum_{i < j} G_{ij} \ln^2 \left(\frac{k_i^c}{k_j^c} \right)$$

Dependence on Kinematics and Jet Radius



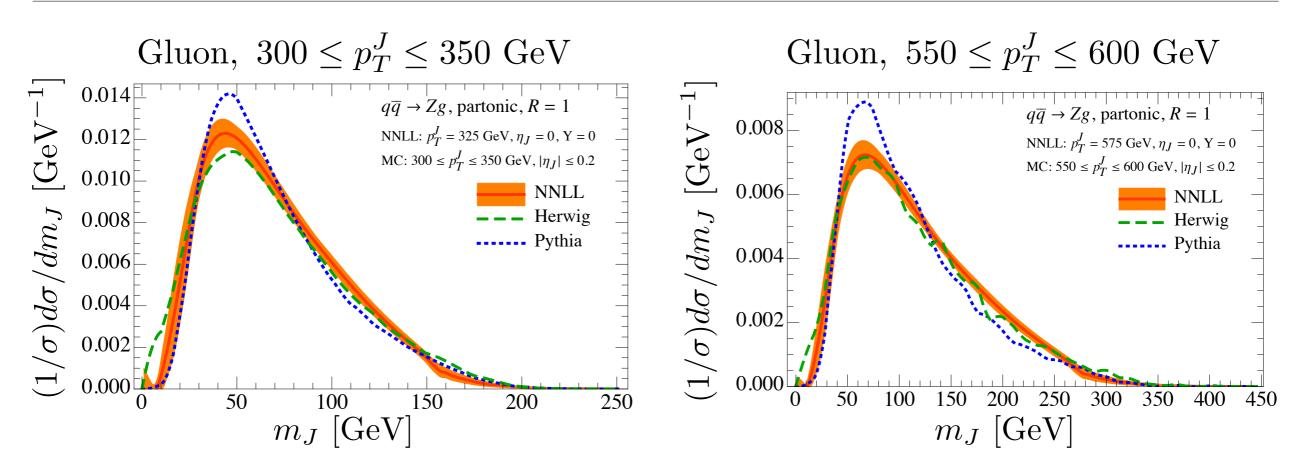
- Calculable dependence on kinematics p_T^J, y_J, Y
- Strong dependence on jet radius since $m_J \lesssim p_T^J R/\sqrt{2}$ (Nonsingular important!)

Comparison to Pythia and Herwig



- Reasonable agreement over a range of kinematics and R
- No clear favorite between Pythia or Herwig
- Big differences for R < 0.5

Comparison to Pythia and Herwig

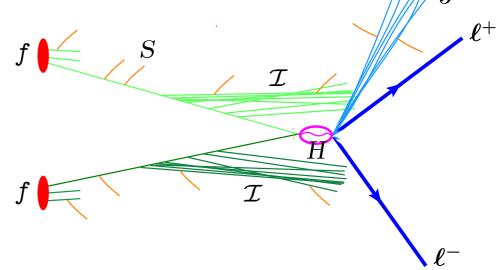


- Reasonable agreement over a range of kinematics and R
- No clear favorite between Pythia or Herwig
- Big differences for R < 0.5

Hadronization of Jets

Tackmann, Stewart, WW (to appear)

Nonperturbative Effects



$$\frac{d\sigma}{dm_J^2} = ff\,\mathcal{I}\,\mathcal{I}\,H \int\! dk_s\,J(m_J^2-2p_T^Jk_s)\,S(k_s)$$

 Jet function Soft function

Soft function describes primary soft radiation:

$$S(k_s) = \langle 0|Y_J^{\dagger}(y_J)Y_{\bar{n}}^{\dagger}Y_n^{\dagger} \,\delta(k_s - \cosh y_J \, n_J \cdot \hat{p}_J) \, Y_n Y_{\bar{n}} Y_J(y_J)|0\rangle$$

Perturbative and nonperturbative soft contribution:

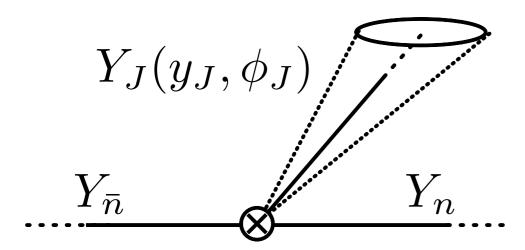
$$S(k_s) = \int dk_s' S_{\text{pert}}(k_s - k_s') F_{\text{NP}}(k_s')$$

(Korchemsky, Sterman; Hoang, Stewart; Ligeti, Stewart, Tackmann)

• Leading NP effect: $m_J^2 \to m_J^2 + 2p_T^J \Omega$, $\Omega = \int dk_s \, k_s \, F_{\rm NP}(k_s)$

Leading Nonperturbative Effect

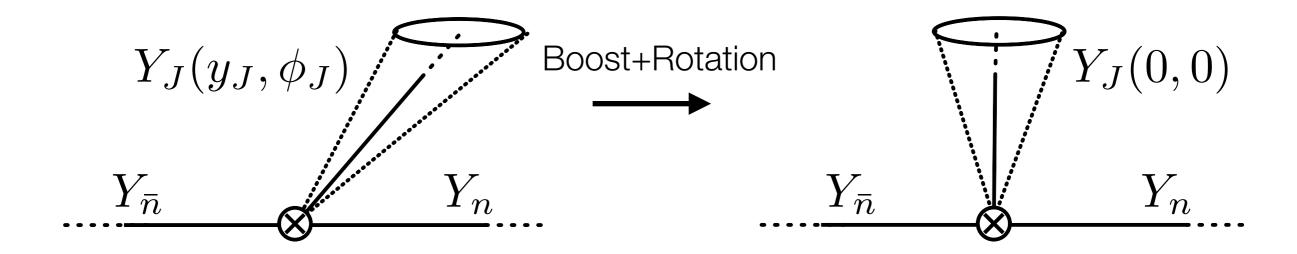
$$\Omega = \langle 0|Y_J^{\dagger}(y_J,\phi_J)Y_{\bar{n}}^{\dagger}Y_n^{\dagger} \cosh y_J \, n_J \cdot \hat{p}_J \, Y_n Y_{\bar{n}} Y_J(y_J,\phi_J)|0\rangle$$



- Ω is independent of p_T^J by definition
- Y's and thus Ω depend on color configuration

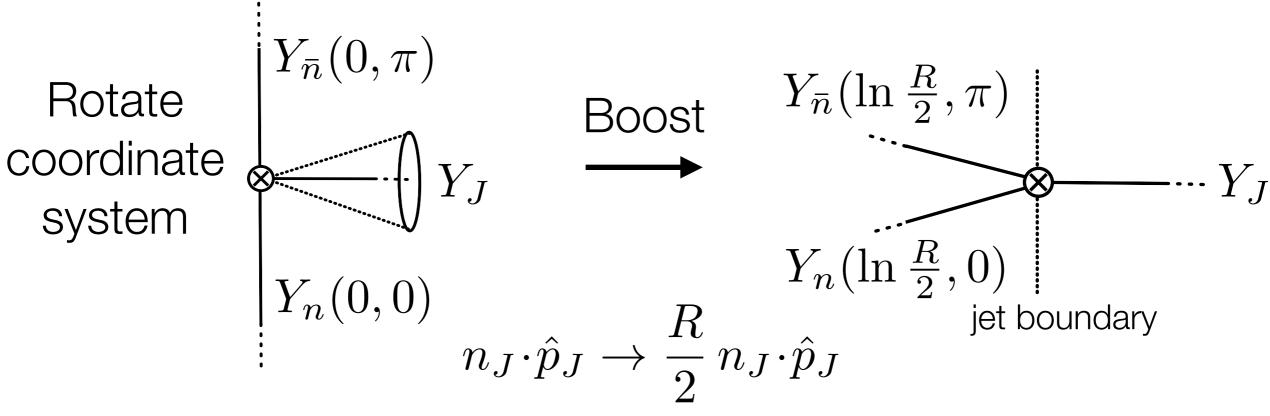
Leading Nonperturbative Effect

$$\Omega = \langle 0|Y_J^{\dagger}(y_J, \phi_J)Y_{\bar{n}}^{\dagger}Y_n^{\dagger} \cosh y_J \, n_J \cdot \hat{p}_J \, Y_n Y_{\bar{n}} Y_J(y_J, \phi_J)|0\rangle$$



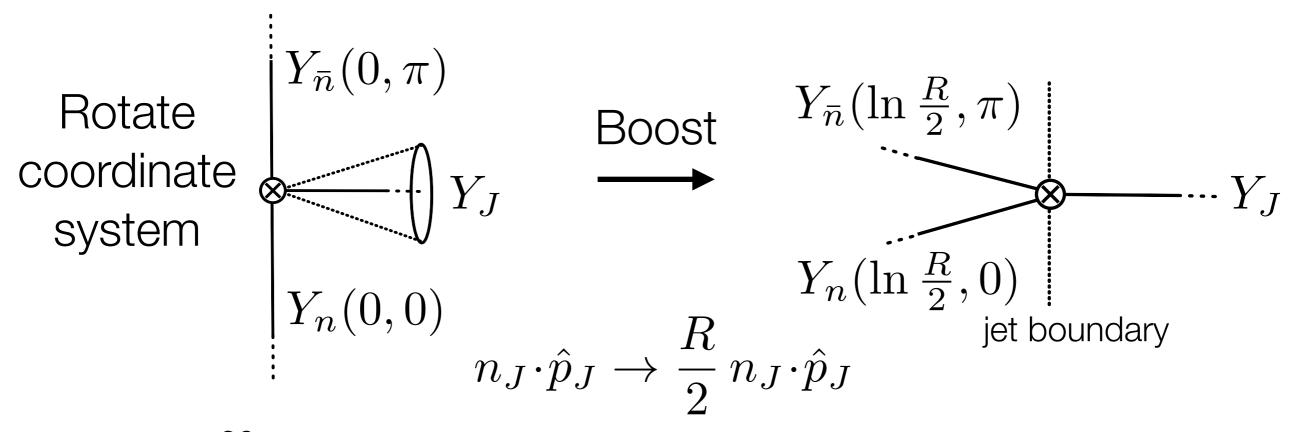
- Ω is independent of p_T^J by definition
- Y's and thus Ω depend on color configuration
- Rotating + boosting shows that Ω is independent of y_J, ϕ_J

Dependence on Jet Radius R



$$\Omega = \frac{R}{2} \int_0^\infty dy \, e^{-y} \langle 0 | Y_J^\dagger Y_{\bar{n}}^\dagger (\ln \tfrac{R}{2}, \pi) Y_n^\dagger (\ln \tfrac{R}{2}, 0) \hat{\mathcal{E}}_\perp(r, y, \phi) (\ldots) | 0 \rangle$$
 energy flow operator

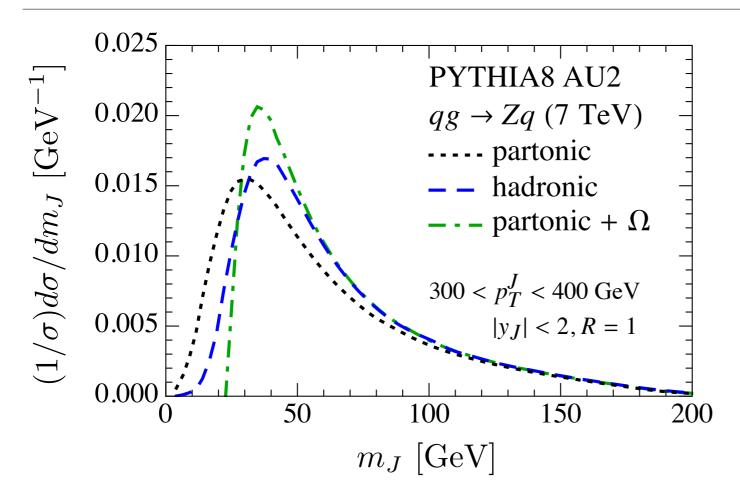
Dependence on Jet Radius R



$$\Omega = \frac{R}{2} \int_0^\infty dy \, e^{-y} \langle 0 | Y_J^\dagger Y_{\bar{n}}^\dagger (\ln \tfrac{R}{2}, \pi) Y_n^\dagger (\ln \tfrac{R}{2}, 0) \hat{\mathcal{E}}_\perp(r, y, \phi) (\ldots) | 0 \rangle$$
 energy flow operator

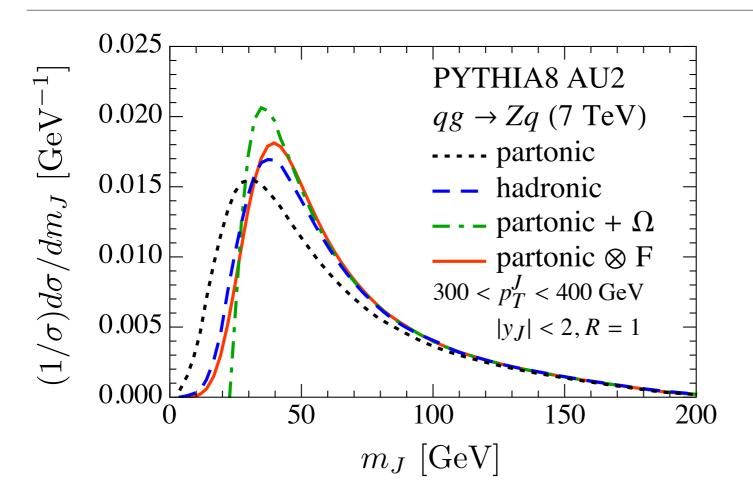
- For $R \ll 1$, the beam Wilson lines fuse and $\Omega = \Omega_0 R/2 + \dots$
- Only odd powers of R arise

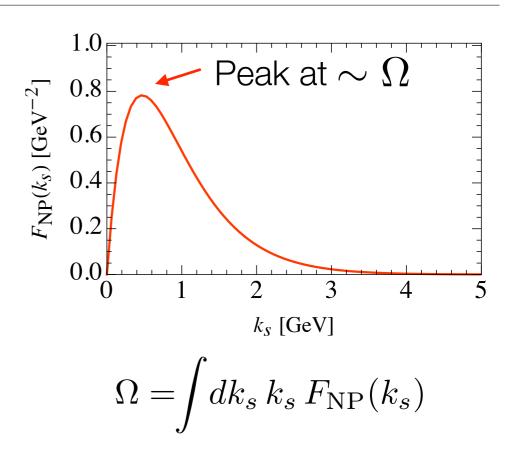
Hadronization in MC



• Hadronization in the tail described by $\; m_J^2
ightarrow m_J^2 + 2 p_T^J \Omega \;$

Hadronization in MC



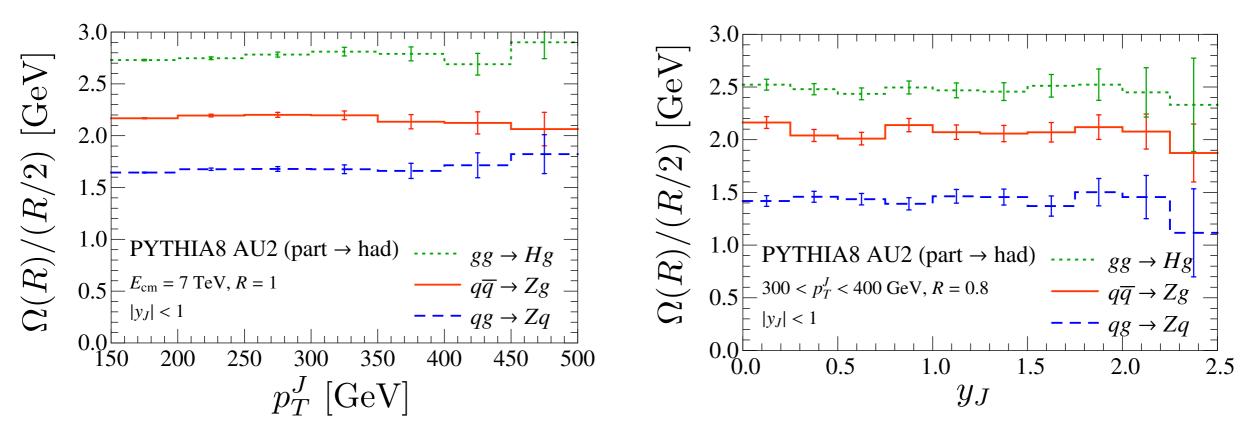


• Hadronization in the tail described by $m_J^2 o m_J^2 + 2p_T^J \Omega$

•
$$\frac{d\sigma}{dm_J^2} \to \int_0^\infty dk_s \, \frac{d\sigma}{dm_J^2} (m_J^2 - 2p_T^J k_s) \, F_{\rm NP}(k_s)$$

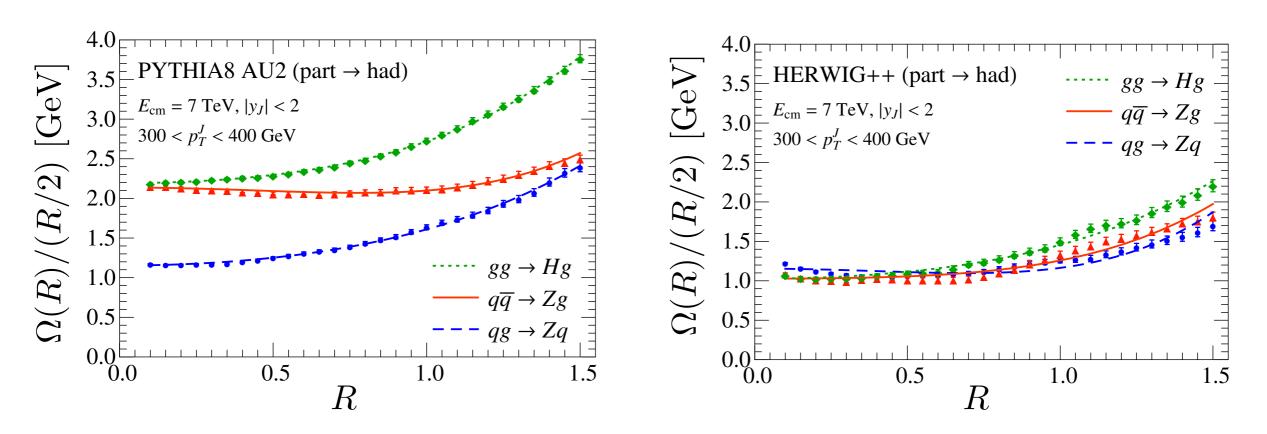
works fairly well for generic $F_{\rm NP}$ with correct Ω

Hadronization is Nonperturbative Soft Effect



• Hadronization in MC is independent of p_T^J, y_J

Hadronization is Nonperturbative Soft Effect



- Hadronization in MC is independent of p_T^J, y_J
- $\Omega = \Omega_0 R/2 + \dots$ for $R \ll 1$
- Ω_0 only depends on quark vs. gluon jet (same in Herwig?)
- The universal Ω_0^q can be extracted from DIS event shapes (DIS Ω_0 : Dasgupta, Salam; Kang, Liu, Mantry, Qiu; Kang, Lee, Stewart)

Conclusions

- Many LHC searches involves jets as signal or background
- Jet substructure provides a new set of tools
 - Boosted objects
 Quark vs. gluon
- Much theoretical work remains to be done
 - Improve Monte Carlo
 Gain insight
- Factorization is key: separating physics at different scales
 - Calculate jet mass and charge
 - Nonperturbative effects for track-based observables
 - Universality of hadronization for jets with $R \ll 1$