Jet Substructure at the LHC

Wouter Waalewijn



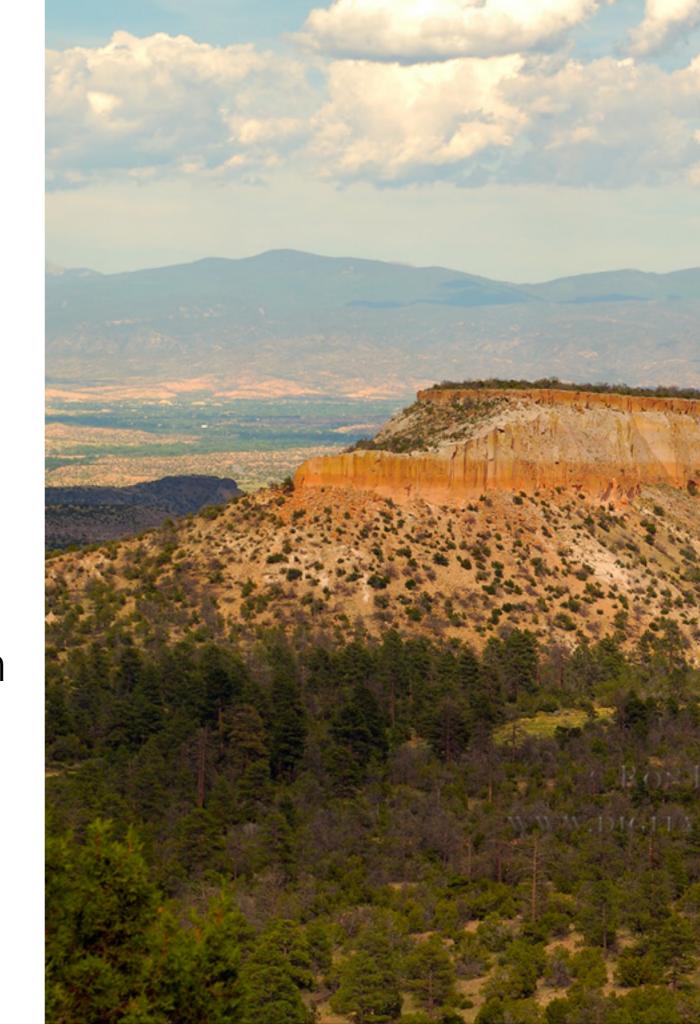




LANL - January 8, 2015

Outline

- Introduction
- Jet Charge
- Jet Mass
- Hadronization of Jets
- Quark/Gluon Discrimination
- Conclusions

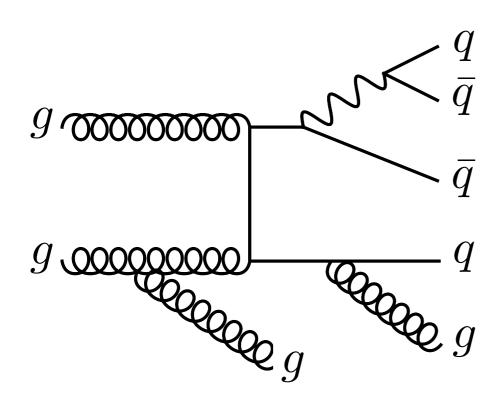


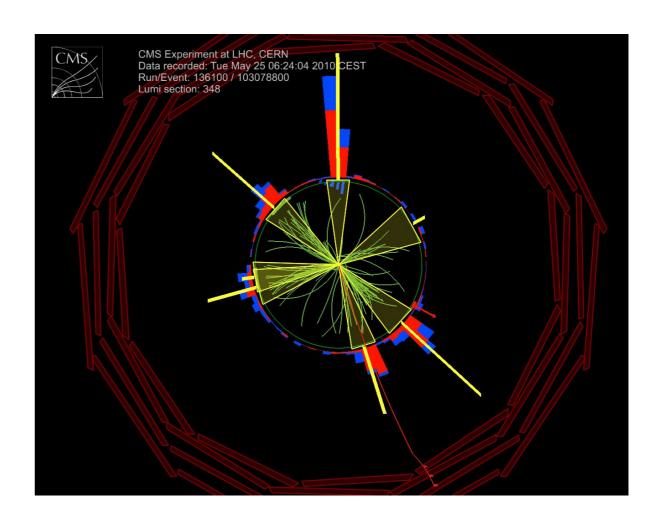
Introduction

What is a Jet?

Energetic quarks and gluons radiate and hadronize

→ Produce jets of hadrons

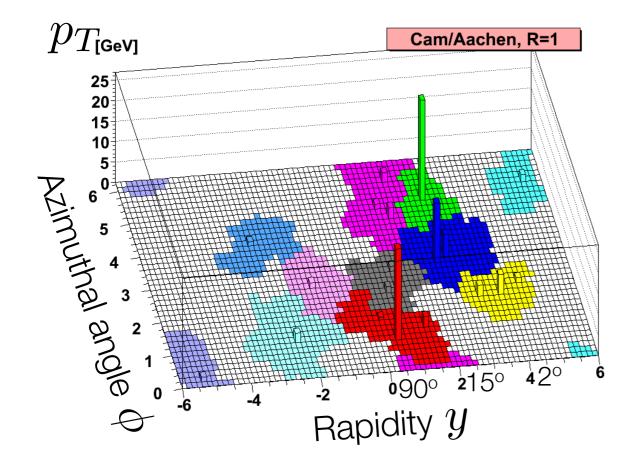




Jet Algorithms

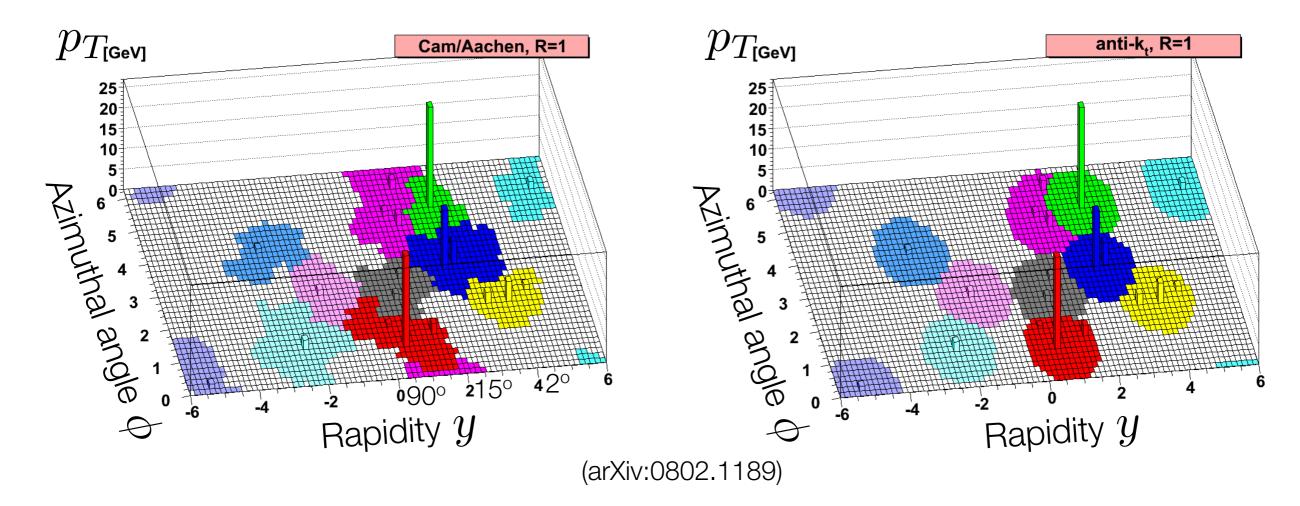
- Repeatedly cluster nearest "particles" $p_i, p_j \rightarrow p_i + p_j$
- Cut off by jet "radius" R

distance =
$$(\Delta y)^2 + (\Delta \phi)^2$$



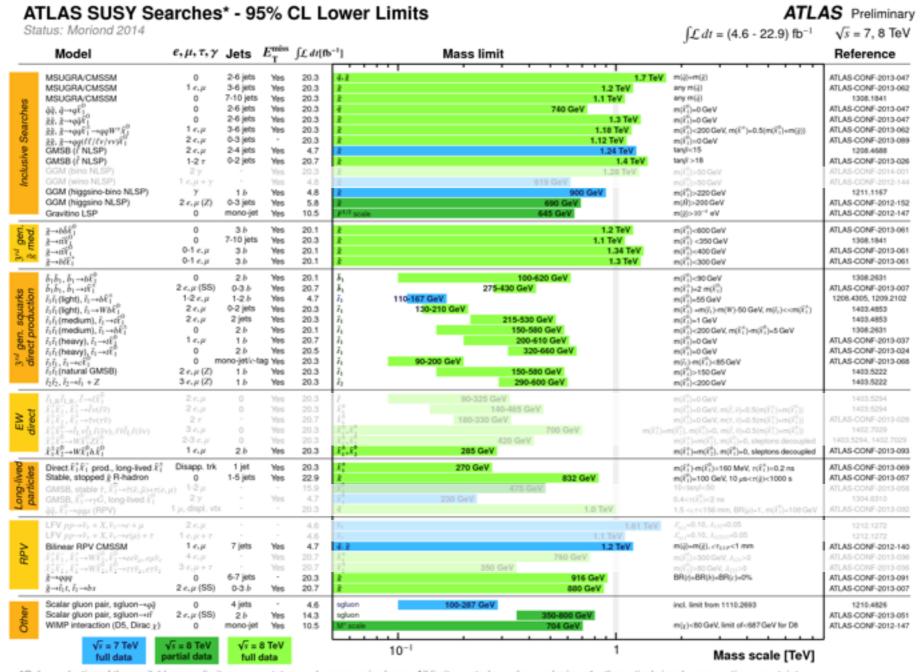
Jet Algorithms

- 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)



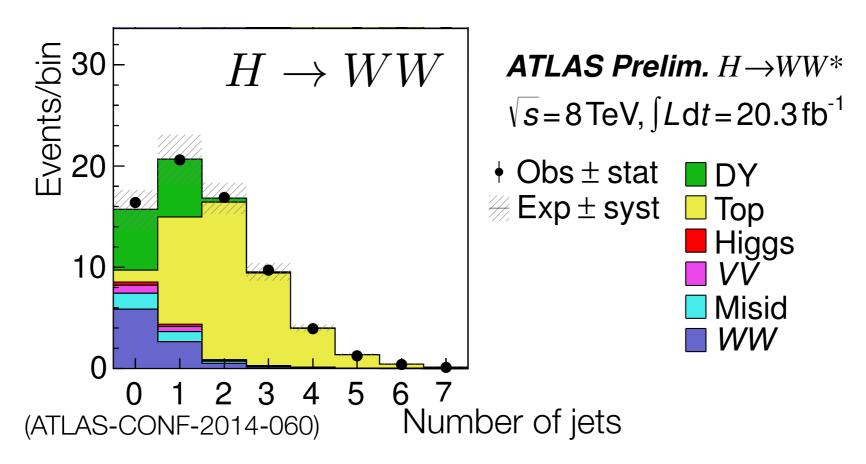
Jets at the LHC

Most measurements involve jets as signal or background



Jet Cross Sections

Bin by jet multiplicity to improve background rejection



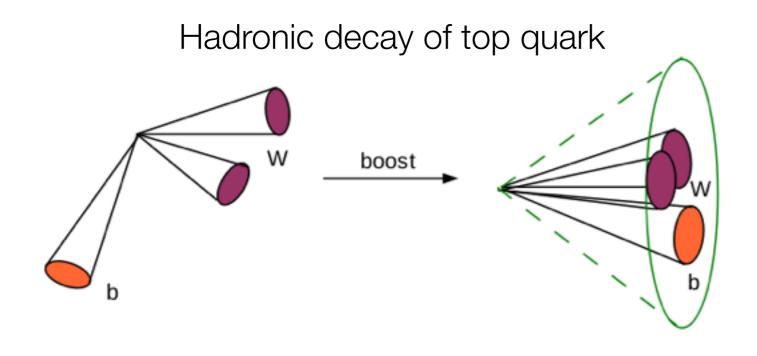
Large logarithms lead to large theory uncertainties

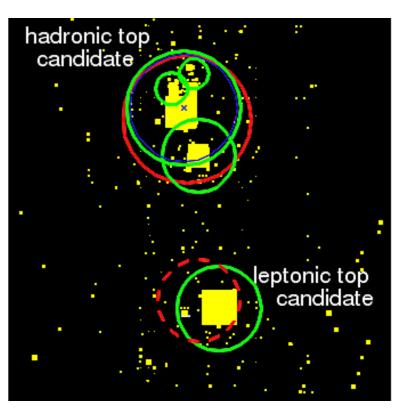
$$\sigma(H+0~{
m jets})\propto 1-rac{6lpha_s}{\pi}\ln^2rac{p_T^{
m cut}}{m_H}+\ldots$$

(Berger, Marcantonini, Stewart, Tackmann, WW; Banfi, Monni, Salam, Zanderighi, Becher, Neubert, Rothen; Stewart, Tackmann, Walsh, Zuberi; Liu, Petriello; Boughezal, Focke, Li, Liu; Jaiswal, Okui, ...)

Jet Substructure for Boosted Objects

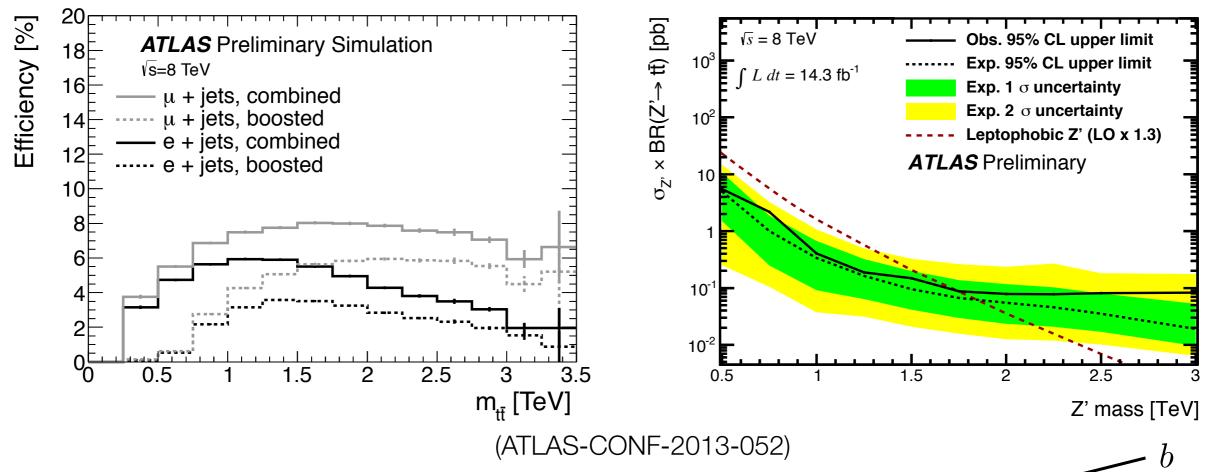
- New heavy particles could produce boosted top, W, Higgs
 - → decay products lie within one "fat" jet
- Distinguish from QCD jets using jet substructure
- Avoids combinatorial background



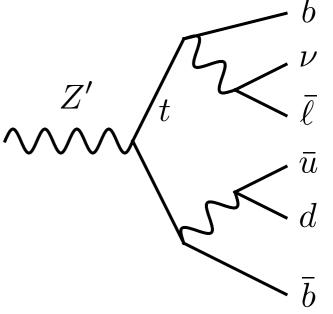


(ATLAS-CONF-2013-052)

Top Tagging in $Z' \to t\bar{t}$



- One leptonic and one hadronic top
- Boosted analysis crucial for large $m_{Z'}$

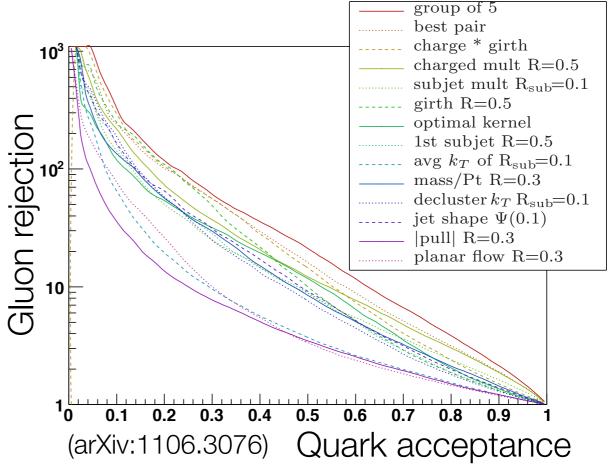


Jet Substructure for Quark/Gluon Discrimination

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

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

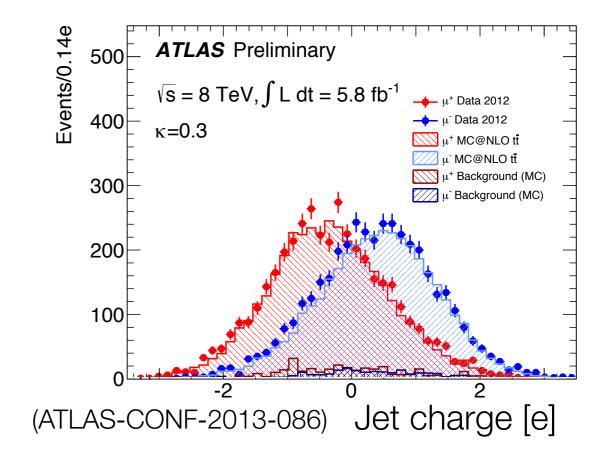
 More variables only give marginal improvement

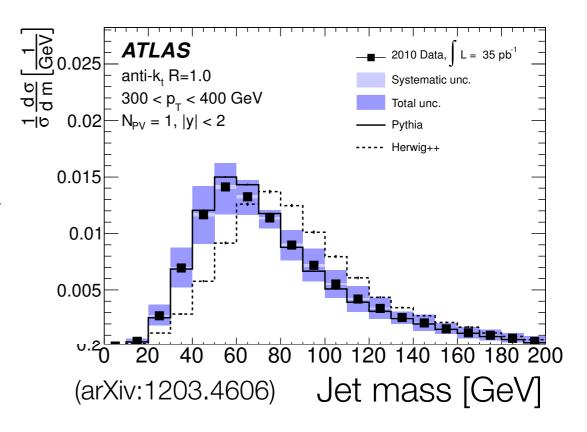


Jet Mass and Charge

Motivation:

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





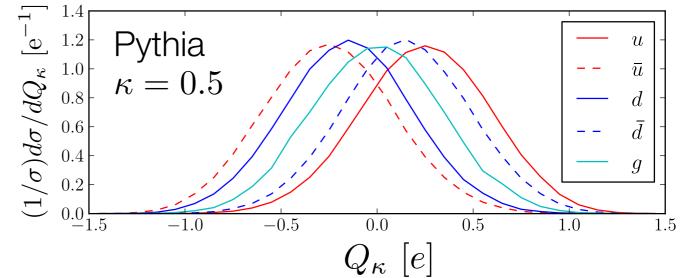
Jet Charge

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

WW (arXiv:1209.3091)

Defining Jet Charge

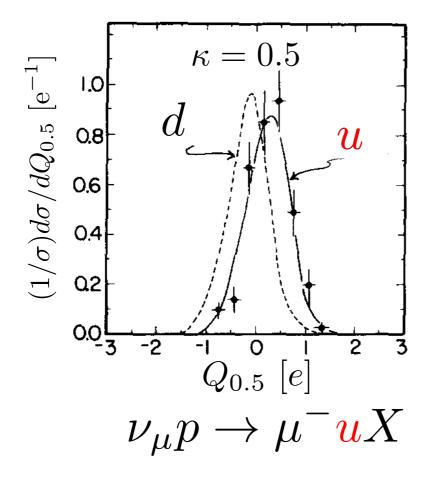
$$Q_{\kappa} = \sum_{i \in \mathrm{jet}} Q_i \Big(rac{p_T^i}{p_T^J} \Big)^{\kappa}$$
 $\mathcal{Q}_{0.8}^{[\frac{1.4}{2}]} \mathcal{Q}_{0.8}^{[\frac{1.4}{2}]}$ Pythia $\kappa = 0.5$ (Feynman, Field)

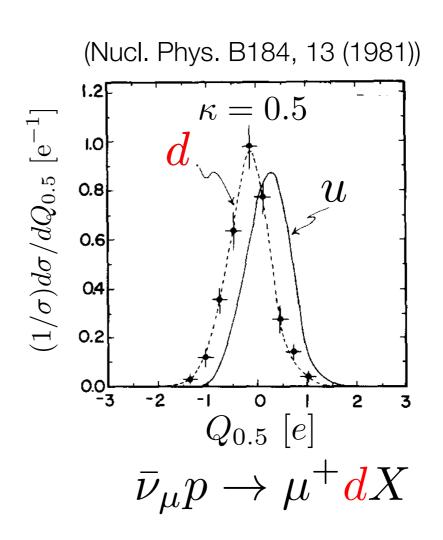


- If κ too small: sensitive to soft hadrons \to contamination
- If κ too large: only sensitive to most energetic hadron
 - → need more statistics

Historical Applications

Test parton model



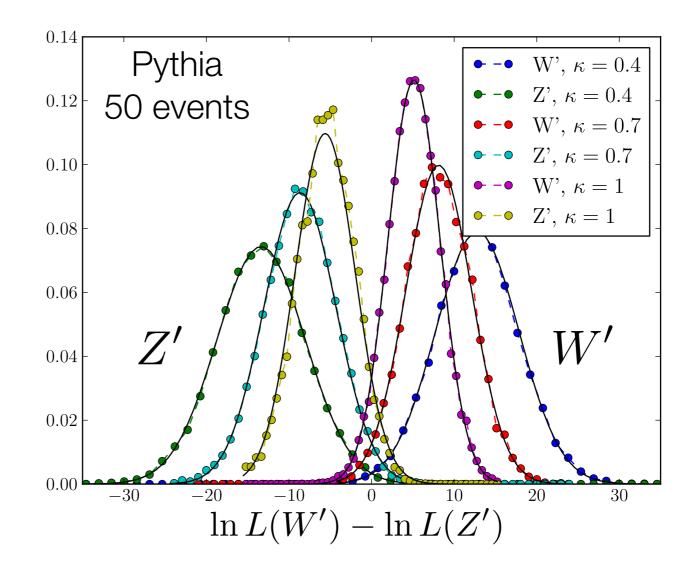


- Jet charge at LEP:
 - Forward-backward charge asymmetry (AMY (1990),...)
 - $B^0 \leftrightarrow \overline{B^0}$ mixing (ALEPH (1992), ...)

Possible LHC application: W' vs. Z'

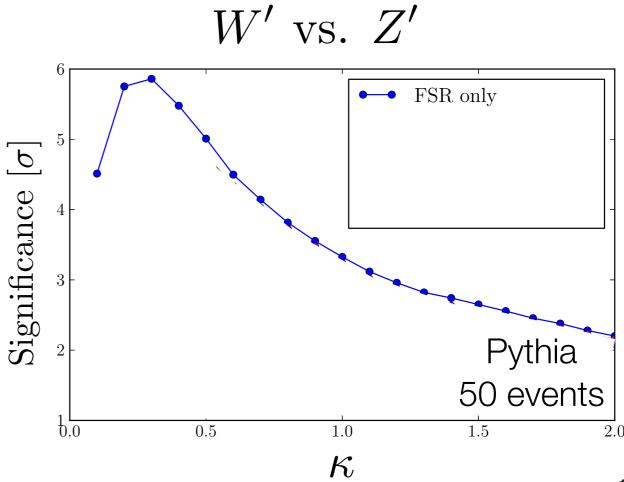
- Hadronically decaying W' or Z' with 1 TeV mass
- 2-dim. likelihood discriminant based on both jet charges

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



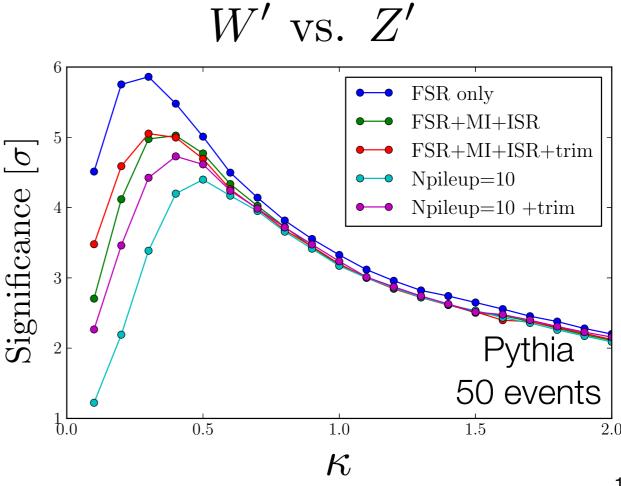
LHC Challenges

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



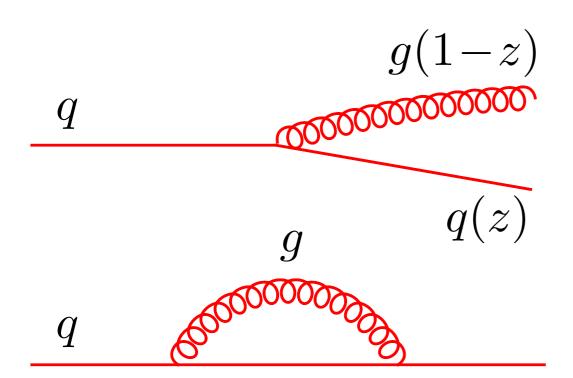
LHC Challenges

- 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



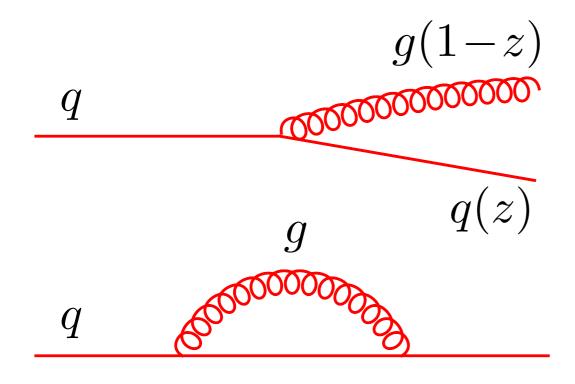
Jet Charge Not IR Safe

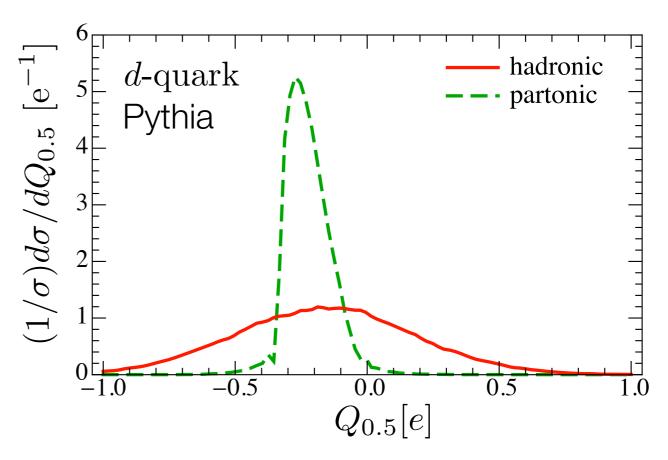
- Consider $q \rightarrow qg$ in collinear limit
- $Q_q z^{\kappa} \neq Q_q$ divergences don't cancel between real/virtual



Jet Charge Not IR Safe

- Consider $q \to qg$ in collinear limit
- $Q_q z^{\kappa} \neq Q_q$ 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

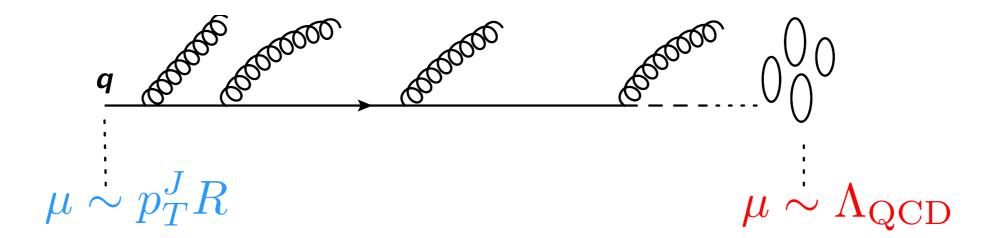
• At LO, weight = fragmentation function $D_q^h(z,\mu \sim p_T^J R)$

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

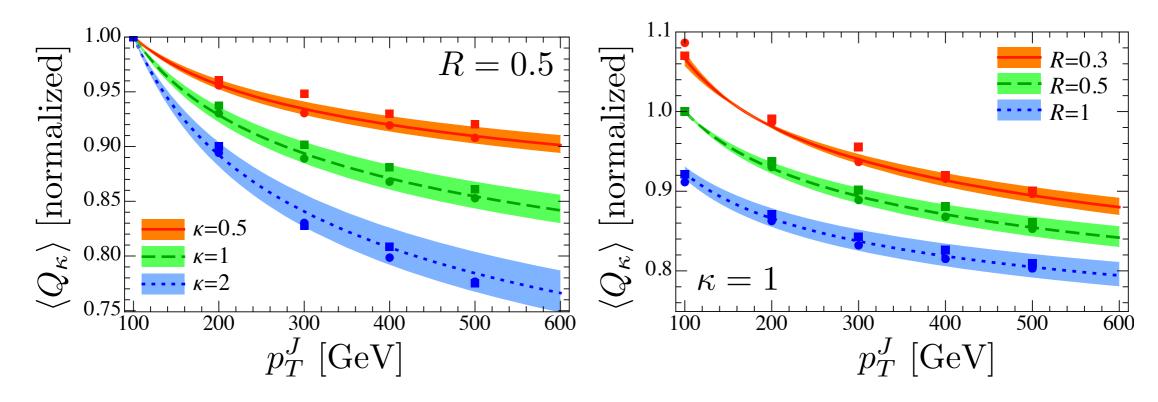


Jet scale

RG Evolution vs. Pythia's Parton Shower

$$\langle Q_{\kappa}(p_T^JR, \mathrm{flavor}) \rangle = \mathrm{perturbative}(\kappa, p_T^JR) \times \mathrm{hadronization}(\kappa, \mathrm{flavor})$$
 perturbative splitting + evolution

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



√ Good agreement

Fragmentation Functions vs. Pythia's Hadronization

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

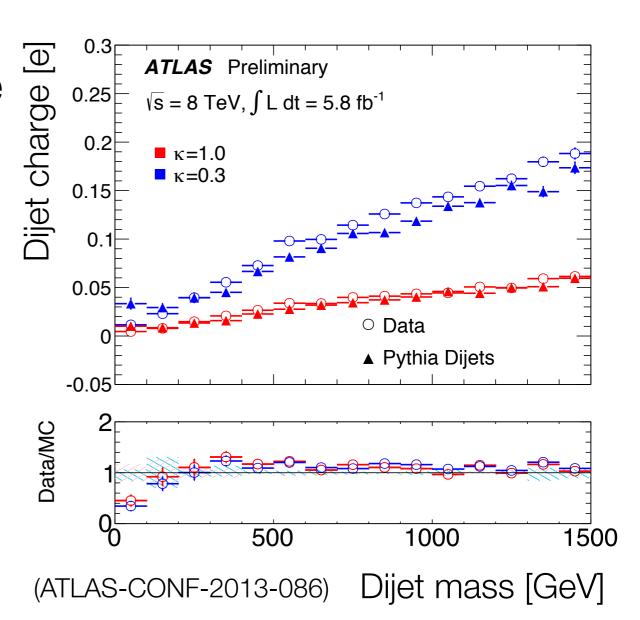
| | $oldsymbol{u}$ -quark | | | <i>d</i> -quark | | |
|----------|-----------------------|-------|-------|-----------------|--------|--------|
| κ | PYTHIA | DSS | AKK08 | PYTHIA | DSS | AKK08 |
| 0.5 | 0.271 | 0.237 | 0.221 | -0.162 | -0.184 | -0.062 |
| 1 | 0.144 | 0.122 | 0.134 | -0.078 | -0.088 |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

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

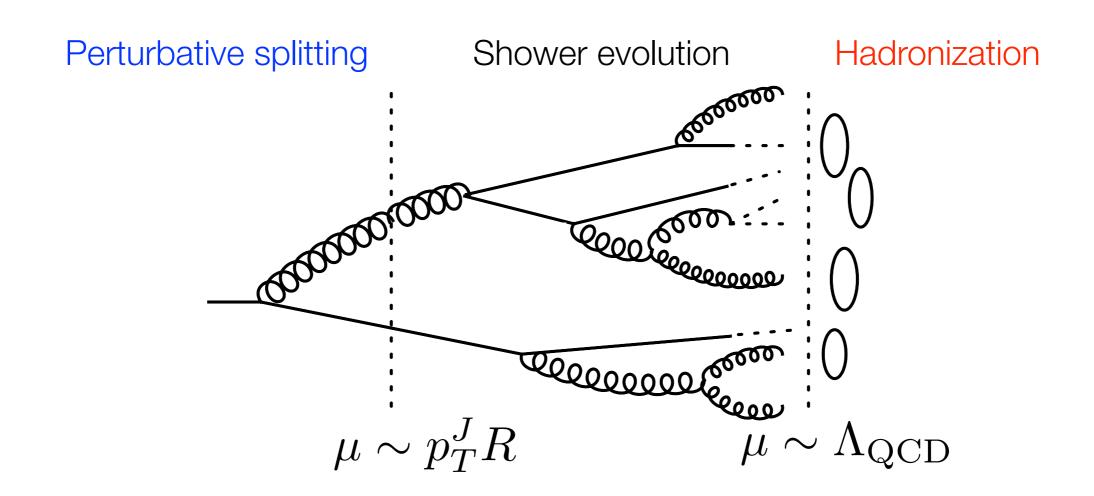
- ✓ Pythia consistent with fragmentation functions
- Large uncertainties as we need $D_q^{h^+}-D_q^{h^-}=D_q^{h^+}-D_{\bar q}^{h^+}$ Most fragmentation data is e^+e^- giving $D_q^{h^+}+D_{\bar q}^{h^+}$

Average Dijet Charge at the LHC

- Depends on proton structure and scattering process
- Pure QCD measurement of valence structure of proton!
- Study of scale violation effect is ongoing

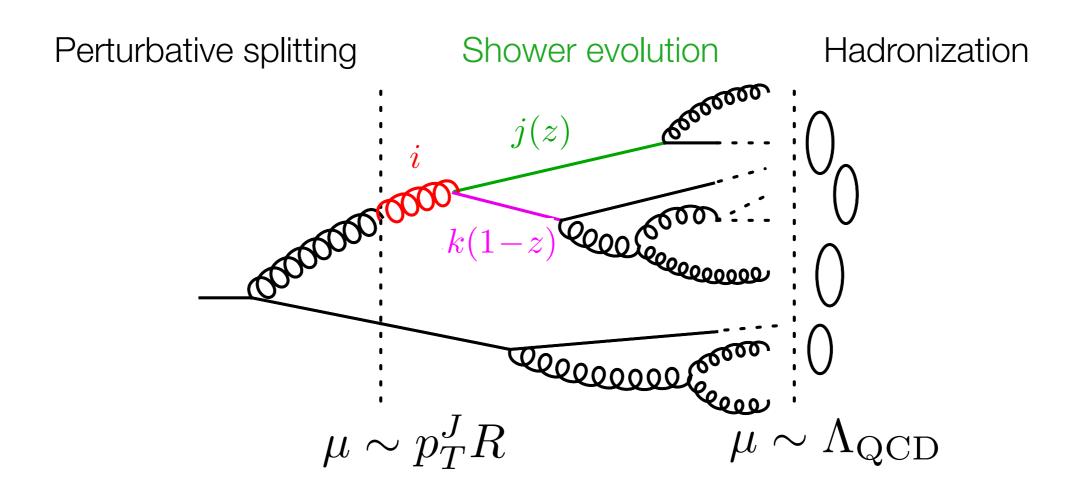


Full Jet Charge Distribution



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

Full Jet Charge Distribution

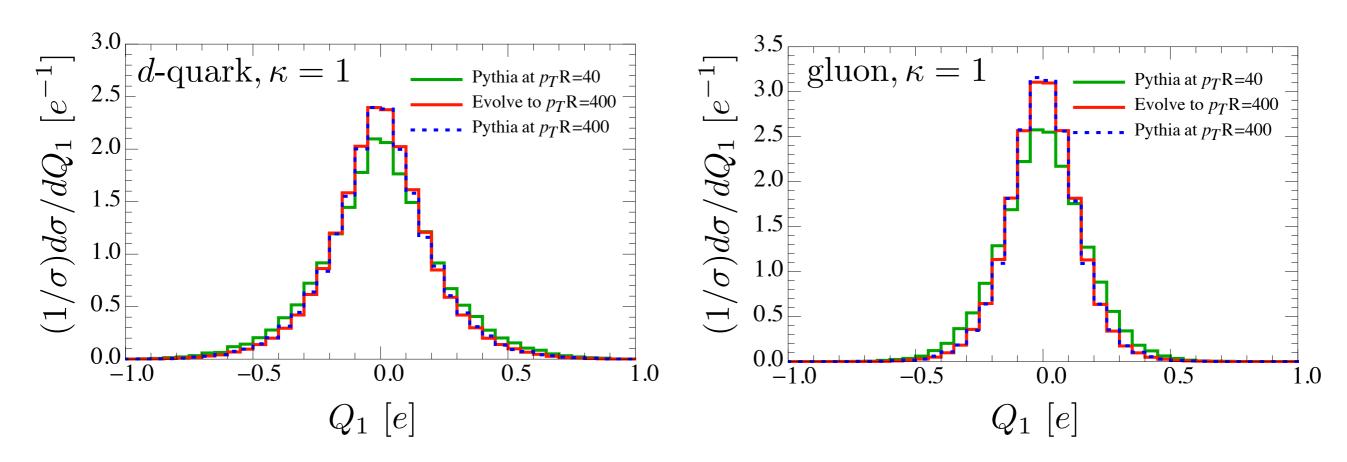


PRGE: Splitting probability Sample over distributions of branches
$$\mu \frac{d}{d\mu} D_i(Q_{\kappa}, \mu) = \sum_j \int dz \frac{\alpha_s}{2\pi} P_{ji}(z) \int dQ_{\kappa}^a D_j(Q_{\kappa}^a, \mu) \int dQ_{\kappa}^b D_k(Q_{\kappa}^b, \mu)$$

$$\times \underbrace{\delta[Q_{\kappa} - z^{\kappa}Q_{\kappa}^a - (1-z)^{\kappa}Q_{\kappa}^b]}_{\text{Charge is (weighted) sum of branches}}$$

RG Evolution vs. Pythia's Parton Shower

- ✓ Use Pythia as input and evolve → good agreement
- Distribution changes more slowly than single hadron distributions (e.g. fragmentation functions)



Jet Mass

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

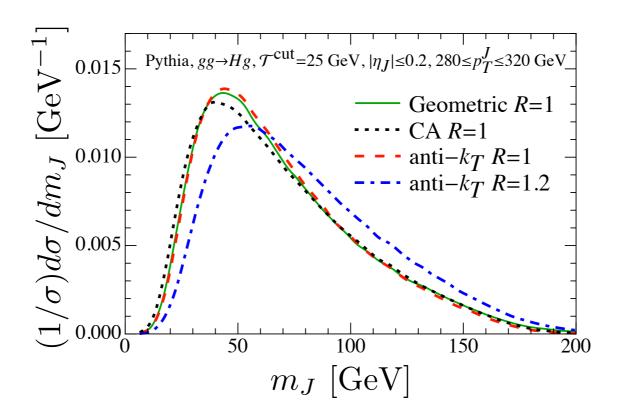
Jet Mass Resummation

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

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

- Need to resum dominant higher-order effects for $m_J^{\mathrm{cut}} \ll p_T^J$
- 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

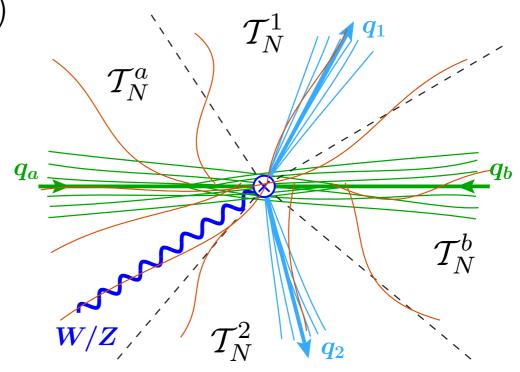
- Reference vectors: $\hat{q}_{a,b}=(1,0,0,\pm 1)$, $\hat{q}_J=(1,\hat{n}_J)/\rho_J$
- $\mathcal{T}_N \to 0$ for N narrow jets, \mathcal{T}_N large for > N jets
- Used as substructure (Thaler, van Tilburg), 1-jettiness in DIS (Kang, Liu, Mantry, Qiu; Kang, Lee, Stewart)

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 N narrow jets, \mathcal{T}_N large for > N jets
- Used as substructure (Thaler, van Tilburg), 1-jettiness in DIS (Kang, Liu, Mantry, Qiu; Kang, Lee, Stewart)
- \mathcal{T}_N splits into contributions from each beam/jet region
- Related to jet mass:

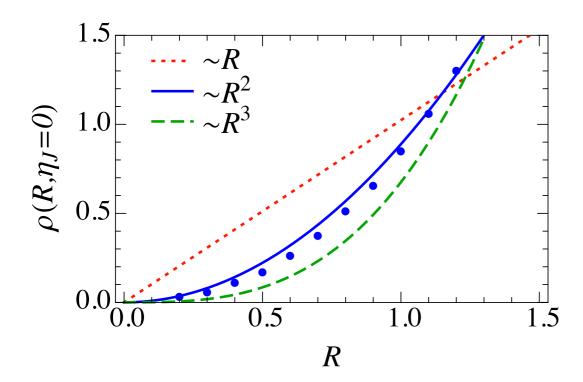
$$m_J^2 = 2\rho_J E_J \mathcal{T}_N^J$$

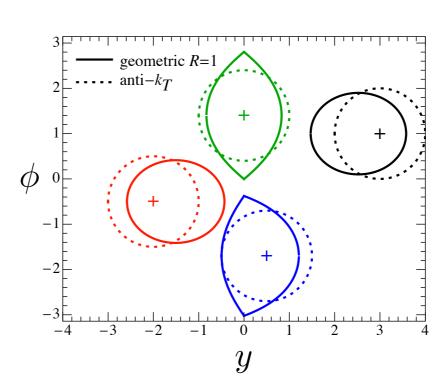


N-Jettiness Parameters

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

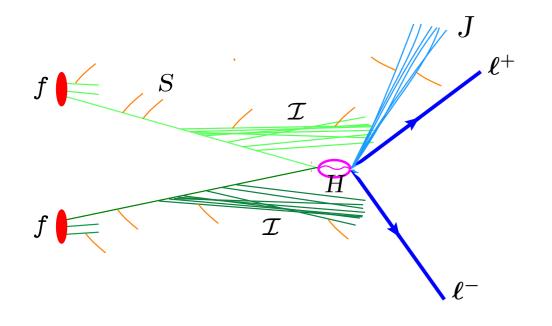
- Reference vectors: $\hat{q}_{a,b} = (1,0,0,\pm 1), \; \hat{q}_{J} = (1,\hat{n}_{J})/\rho_{J}$
- \hat{n}_J by minimizing \mathcal{T}_N or from jet alg. (same for $\mathcal{T}_N \to 0$)
- Choose $ho_J =
 ho(R, \eta_J)$ to match jet area of anti- k_T





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.

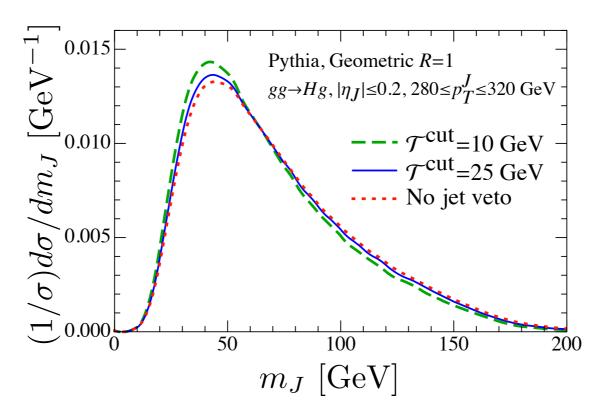
Three-loop cusp: Korchemsky, Radyushkin; Moch, Vermaseren, Vogt, Two-loop non-cusp known from: Kramer, Lampe; Harlander; Aybat, Dixon, Sterman; Becher, Neubert; Becher, Schwartz; Stewart, Tackmann, WW

Normalization

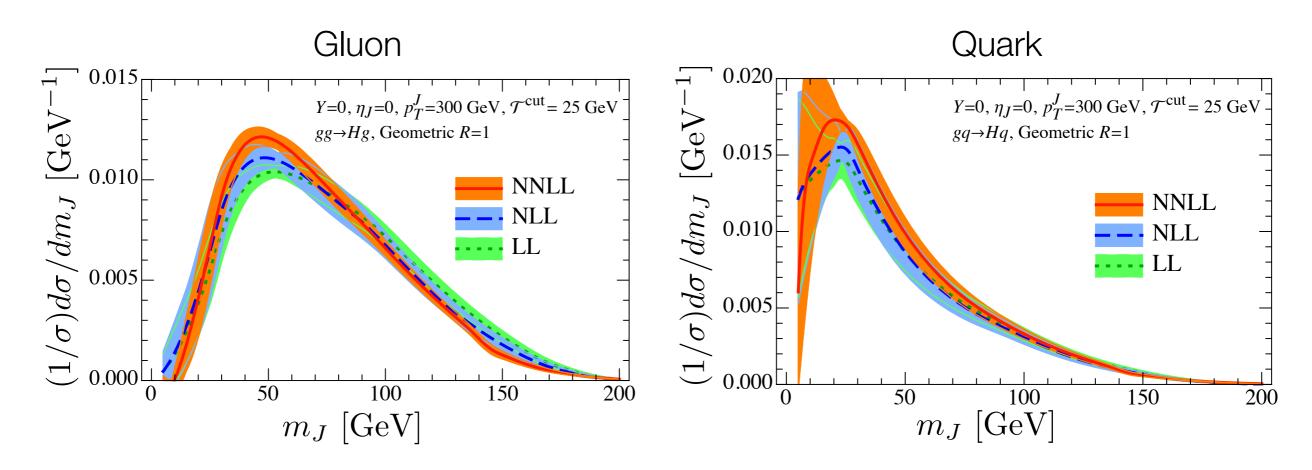
- We are required to veto additional jets through $\mathcal{T}_1^a, \mathcal{T}_1^b$
- Normalizing the spectrum removes this dependence:

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

Experimental results are also normalized

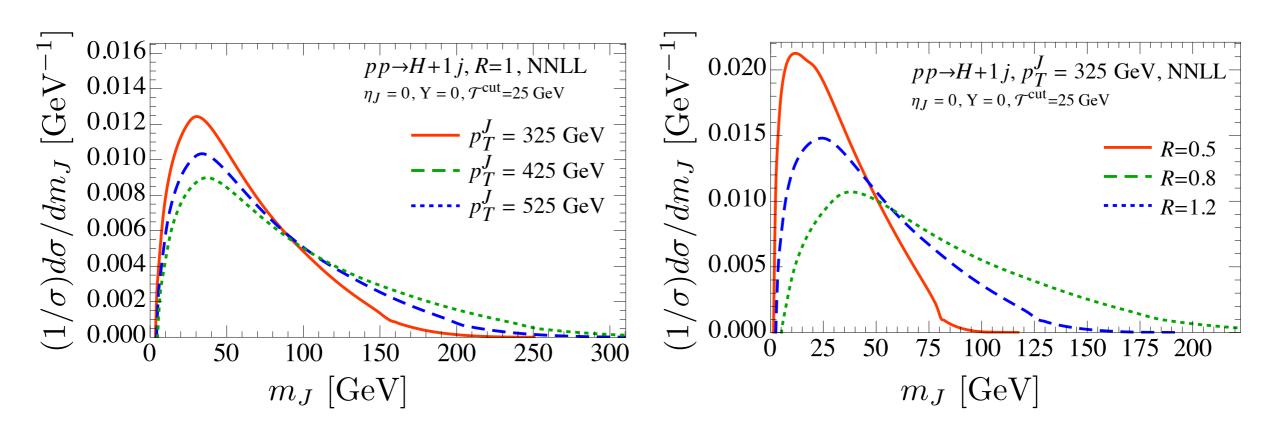


Perturbative Convergence



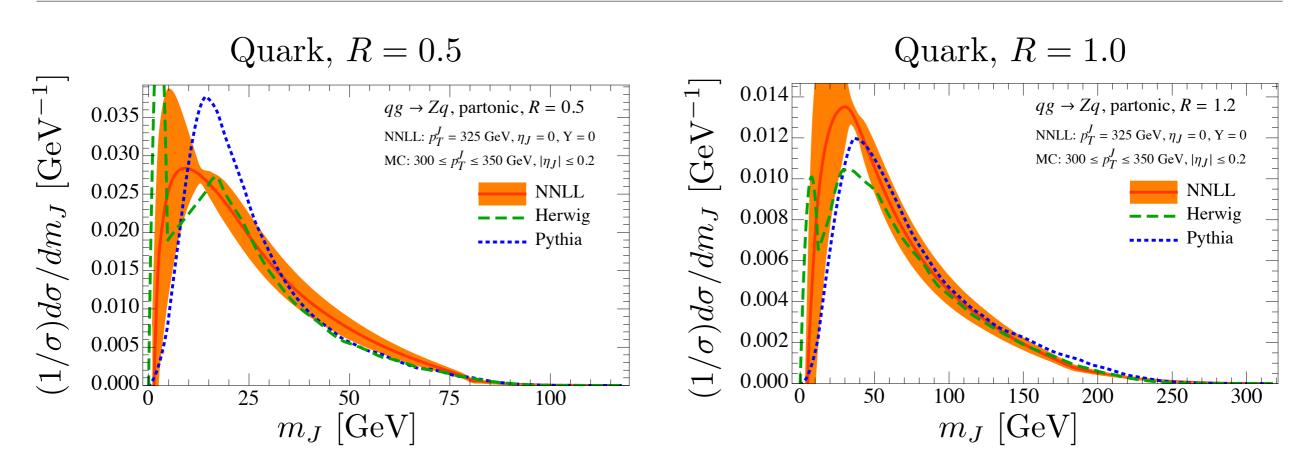
- We consider $gg \to Hg$ and $gq \to Hq$ (proxies for gluon and quark jets)
- √ Good agreement between LL, NLL, NNLL

Dependence on Kinematics and Jet Radius



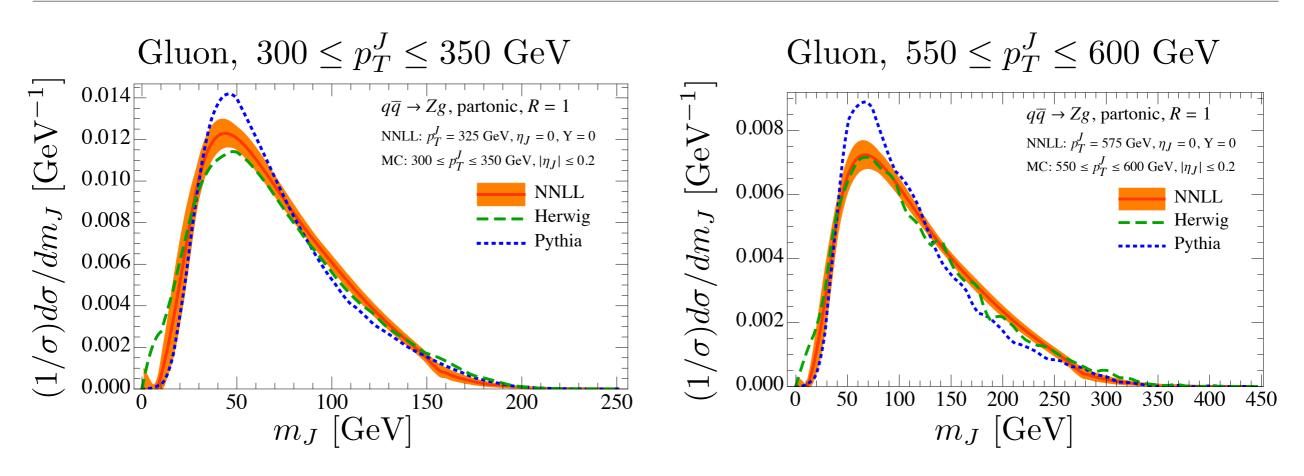
- Calculable dependence on kinematics p_T^J, y_J, Y
- Strong dependence on jet radius since $m_J \lesssim p_T^J R/\sqrt{2}$ (Nonsingular important!)

Comparison to Pythia and Herwig



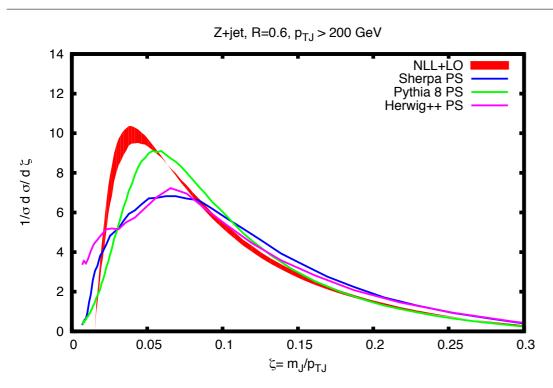
- \checkmark Reasonable agreement over a range of kinematics and R
- No clear favorite between Pythia or Herwig
- Big differences for R < 0.5

Comparison to Pythia and Herwig



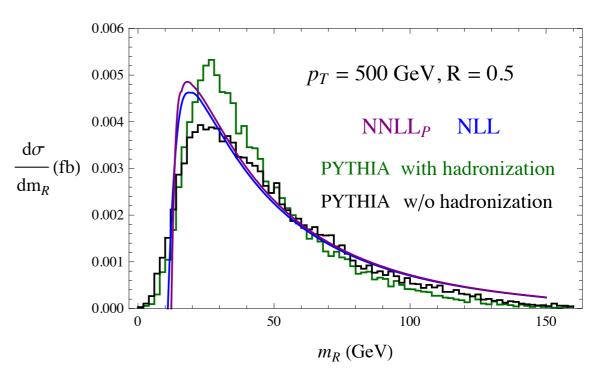
- Reasonable agreement over a range of kinematics and R
- No clear favorite between Pythia or Herwig
- Big differences for R < 0.5

Other Jet Mass Calculations



Dasgupta et al. (arXiv:1207.1640)

- Z+jet and dijets
- NLL+NLO



Chien et al. (arXiv:1208.0010)

- γ +jet
- NNLL threshold resum.

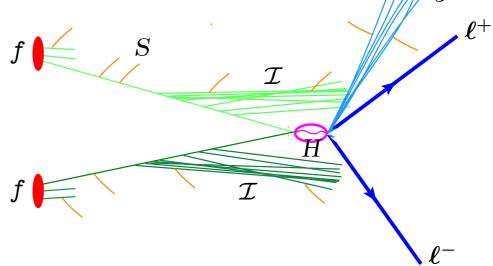
Key differences: • jet algorithm

no jet veto → large nonglobal logarithms

Hadronization of Jets

Tackmann, Stewart, WW (arXiv:1405.6722)

Factorization for Jet Mass



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

 Jet function Soft function

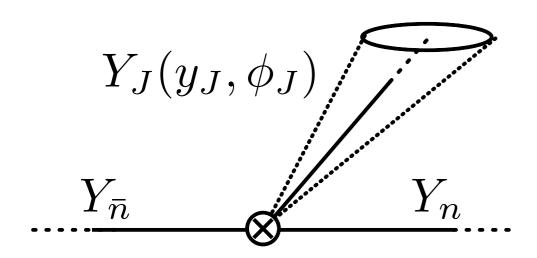
Soft function describes soft radiation:

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

- Color indices on Wilson lines are not written out
- Perturbative and nonperturbative contribution:

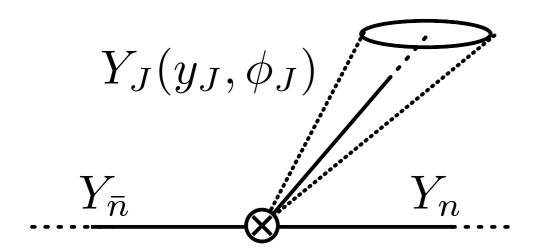
$$S(k_s) = \int dk_s' \, S_{
m pert}(k_s-k_s') F_{
m NP}(k_s') \qquad k_s' \sim \Lambda_{
m QCD}$$
 (Korchemsky, Sterman; Hoang, Stewart; Ligeti, Stewart, Tackmann)

Leading Nonperturbative Effect Ω



- Expanding $F_{\rm NP}(k_s) = \delta(k_s) \Omega \, \delta'(k_s) + \dots$ $\Omega = \langle 0 | Y_J^\dagger(y_J, \phi_J) Y_{\bar{n}}^\dagger Y_n^\dagger \, \cosh y_J \, n_J \cdot \hat{p}_J \, Y_n Y_{\bar{n}} Y_J(y_J, \phi_J) | 0 \rangle$
- Shifts jet mass spectrum $m_J^2 \to m_J^2 + 2p_T^J \Omega$ (valid in tail of distribution)
- Ω is universal for e^+e^- event shapes. (Dokshitzer, Webber; Akhoury, Zakharov; Lee, Sterman; Mateu, Stewart, Thaler) How is this affected by jets?

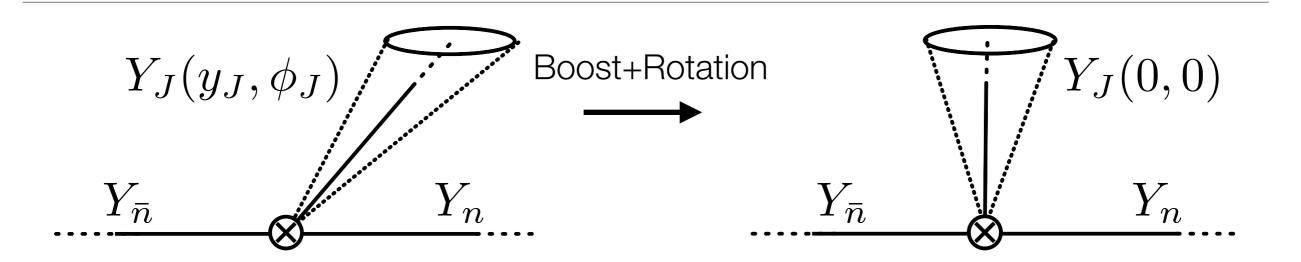
Properties of Ω



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

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

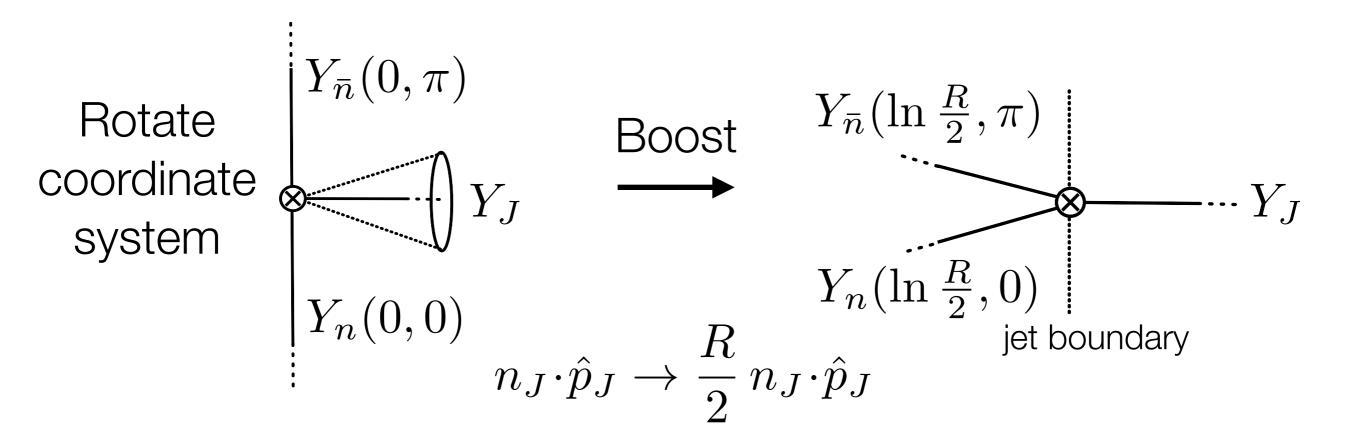
Properties of Ω



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

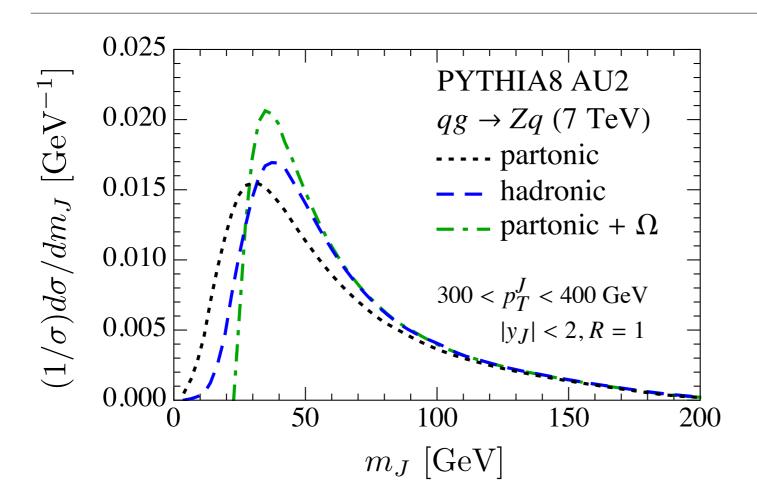
- Ω is independent of p_T^J by definition
- Y's and thus Ω depend on color configuration
- Rotating + boosting shows that Ω is independent of y_J, ϕ_J

Dependence of Ω on Jet Radius R



- For $R\ll 1$, the beam Wilson lines fuse and $\Omega=rac{R}{2}\,\Omega_0+\dots$
- Ω_0 only depends on quark vs. gluon, equal to $\Omega_{\rm DIS}$ (for q) ($\Omega_{\rm DIS}$: Dasgupta, Salam; Kang, Liu, Mantry, Qiu; Kang, Lee, Stewart)
- Only odd powers of R arise

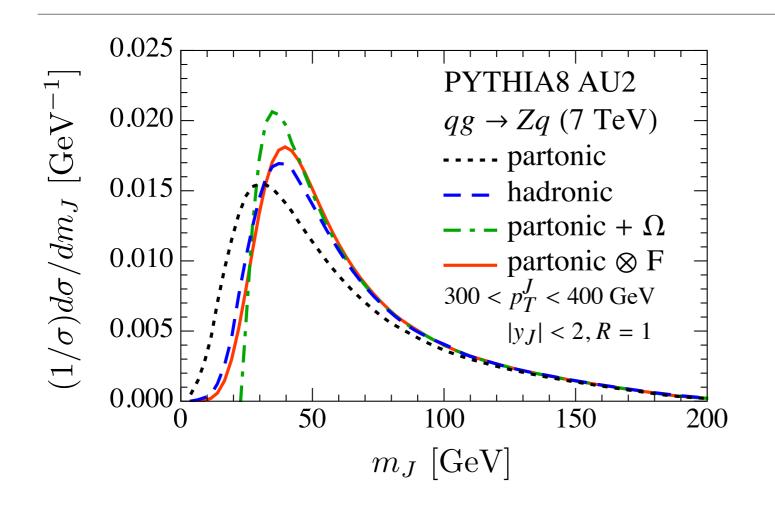
Hadronization captured by Ω

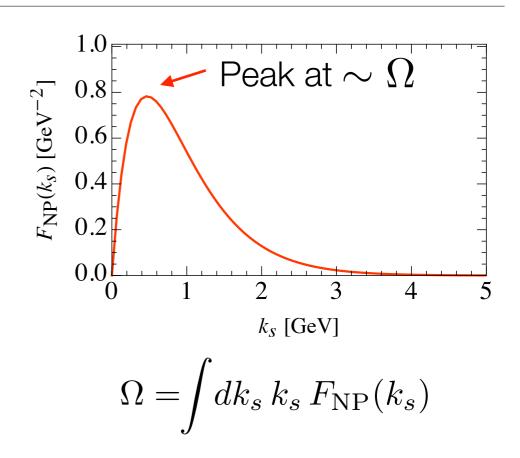


Agrees with factorization predictions:

✓ Hadronization in the tail satisfies $m_J^2 \to m_J^2 + 2p_T^J \Omega$

Hadronization captured by Ω

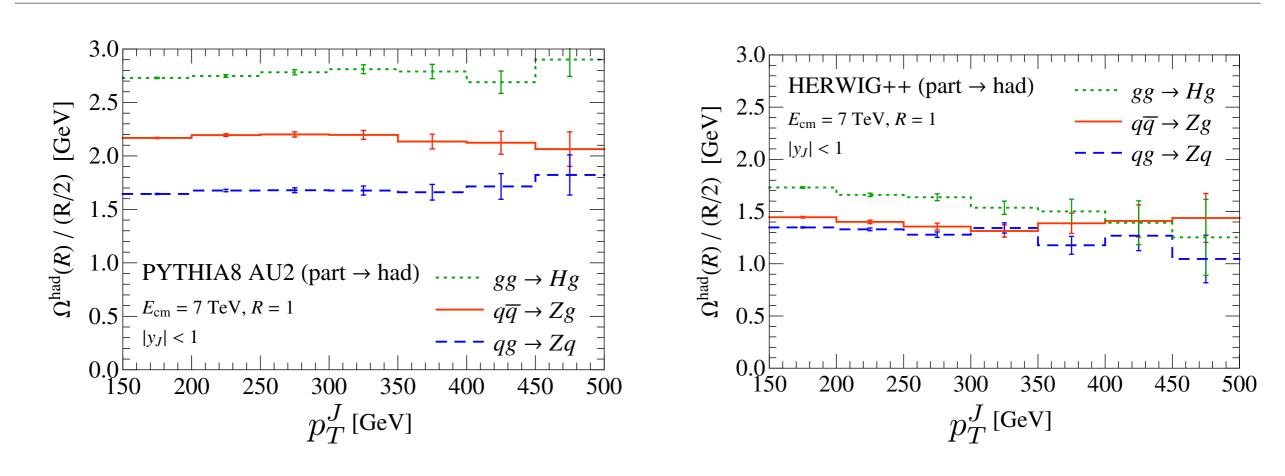




Agrees with factorization predictions:

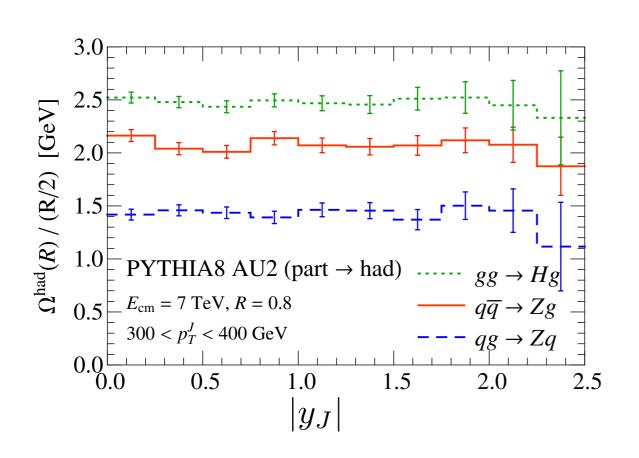
- ✓ Hadronization in the tail satisfies $m_J^2 \to m_J^2 + 2p_T^J \Omega$
- \checkmark More general: $\frac{d\sigma}{dm_J^2} \to \int_0^\infty dk_s \, \frac{d\sigma}{dm_J^2} (m_J^2 2p_T^J k_s) \, F_{\rm NP}(k_s)$

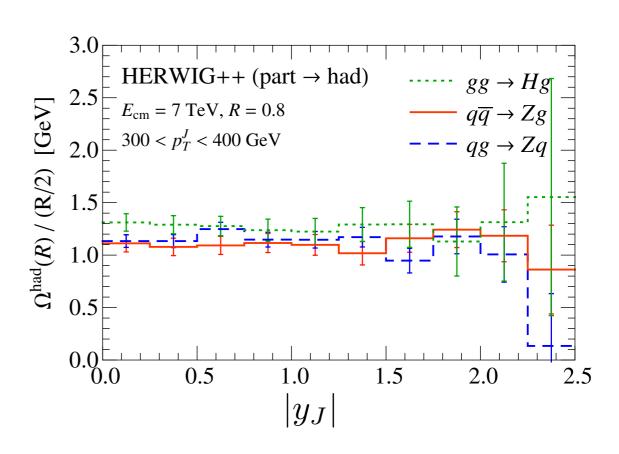
Hadronization dependence on p_T^J



✓ Agrees with factorization predictions

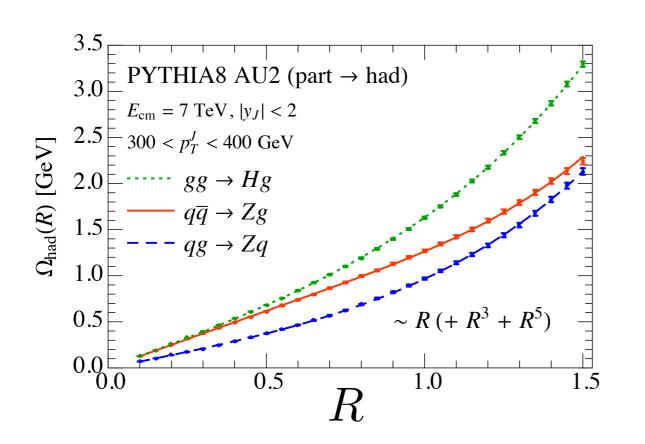
Hadronization dependence on y_J

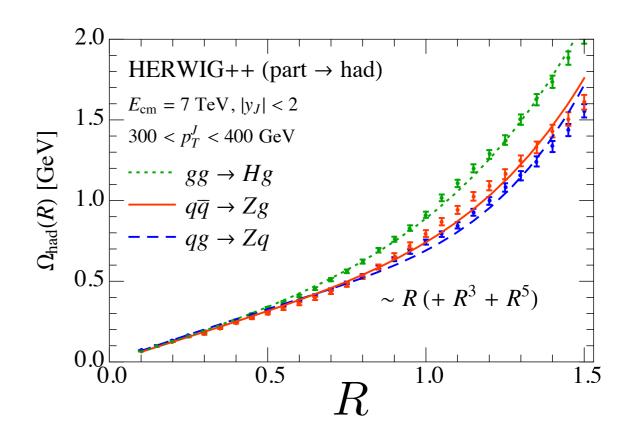




✓ Agrees with factorization predictions

Hadronization dependence on R





- ✓ Linear R coefficient Ω_0 only depends on quark vs. gluon
- ? Quark and gluon jets much more similar in Herwig
- Better fit to odd powers of R in Pythia

Quark/Gluon Discrimination

Larkoski, Thaler, WW (arXiv:1408.3122)

Mutual Information

$$I(A;B) = \int da \, db \, p(a,b) \log_2 \frac{p(a,b)}{p(a)p(b)} \quad \text{A} \qquad \qquad \text{Number of bits of shared information}$$

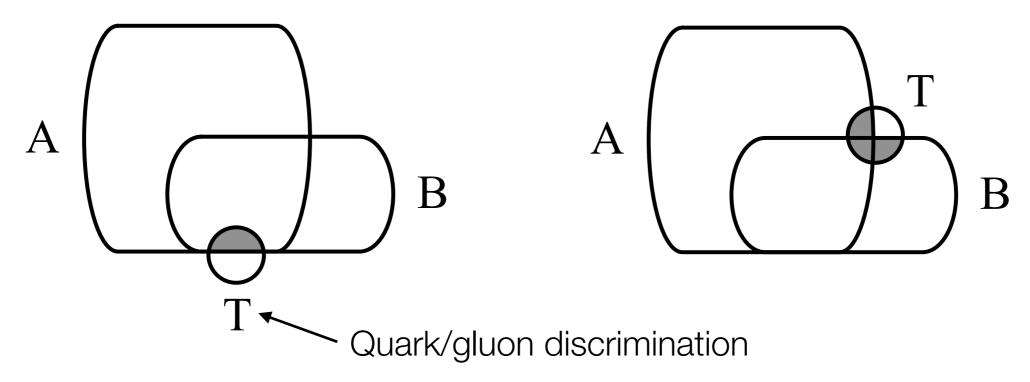
Can directly be calculated from double diff. cross section

$$p(a,b) = \frac{1}{\sigma} \frac{d^2 \sigma}{da \, db}$$

Quark/gluon discrimination is one bit of information

Discrimination Power

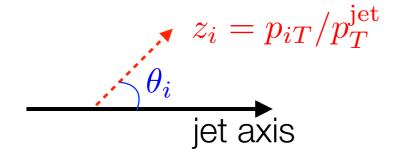
Redundant variables: Complementary variables:

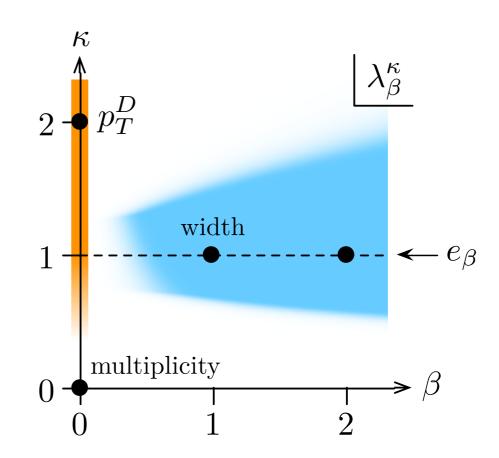


- I(A;B): same correlations
- I(T;A) and I(T;B): same individual discrimination power
- I(T; A, B): different joint discrimination power

Generalized Angularities

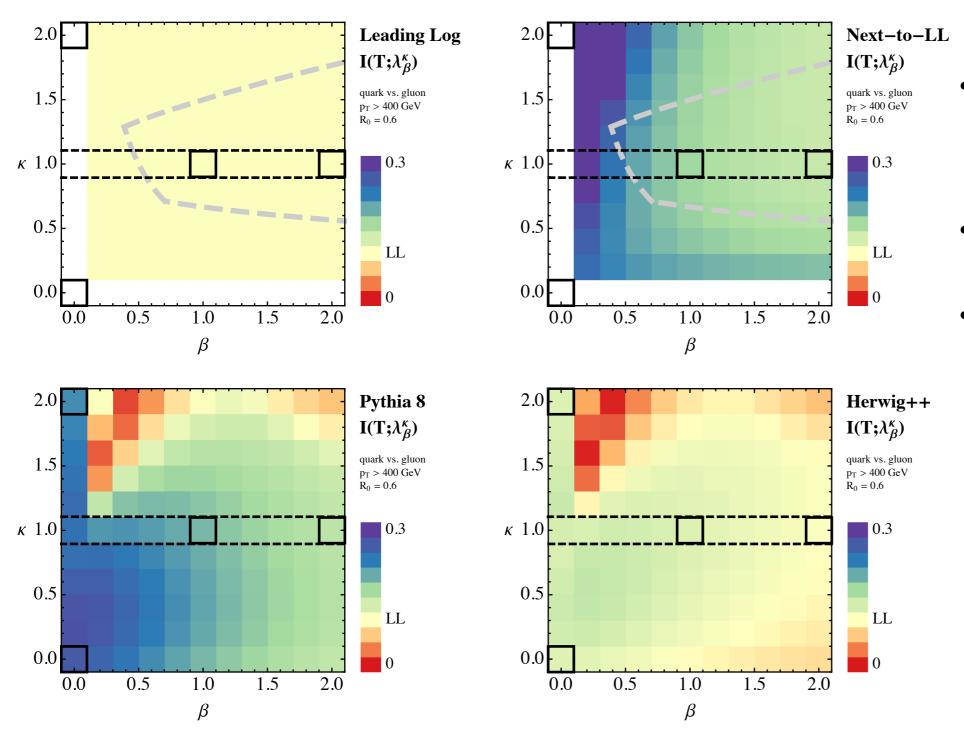
$$\lambda_{\beta}^{\kappa} = \sum_{i \in \text{jet}} z_i^{\kappa} \left(\frac{\theta_i}{R}\right)^{\beta}$$





- $\kappa = 1$: IR safe, angularities (Berger, Kucs, Sterman)
- $\beta = 0$: very IR unsafe, similar to jet charge
- · blue: a bit IR unsafe, one nonpert. parameter at NLL

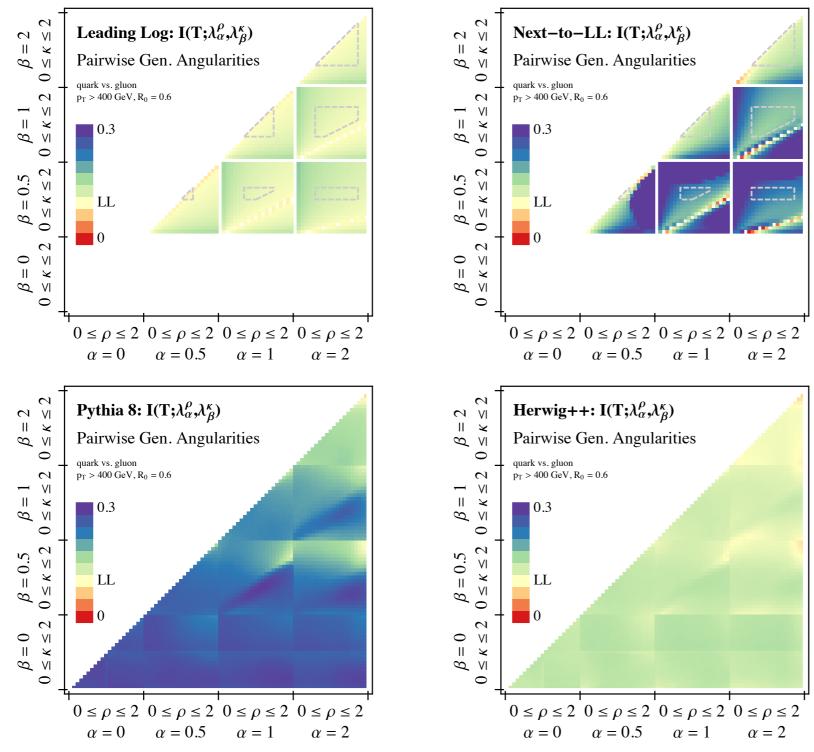
Quark/Gluon Discrimination with λ_{β}^{κ}



- (N)LL valid in grey bounds
- LL is constant
- Significant differences

Calculation uses arXiv:1306.6630 (Chang, Procura, Thaler, WW)

Quark/Gluon Discrimination with $\lambda_{\alpha}^{\rho}, \lambda_{\beta}^{\kappa}$



- (N)LL valid in grey bounds
- LL not const.
- Significant differences

Calculation uses arXiv:1401.4458 (Larkoski, Moult, Neill)

Conclusions

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