# GROOMED JET OBSERVABLES AND CLUSTERING EFFECTS

Kai Yan Harvard University

based on arXiv:1603.06375 with Chris Frye, Andrew Larkoski, Matthew Schwartz

DESY 03/22/2016

#### **GOAL OF THIS TALK**

Theoretical aspects of soft-drop groomed jet substructure:

- ➤ All-order factorization formula for soft-drop groomed observable
- ➤ NNLO analytic calculation that allows complete NNLL resummation

#### **MOTIVATION**

precision comparison between experiment and data due to contamination from UE/PU



filtering/mass drop: Butterworth, Davison, Rubin, Salam 0802.2470

trimming: Krohn, Thaler, Wang 0912.1342

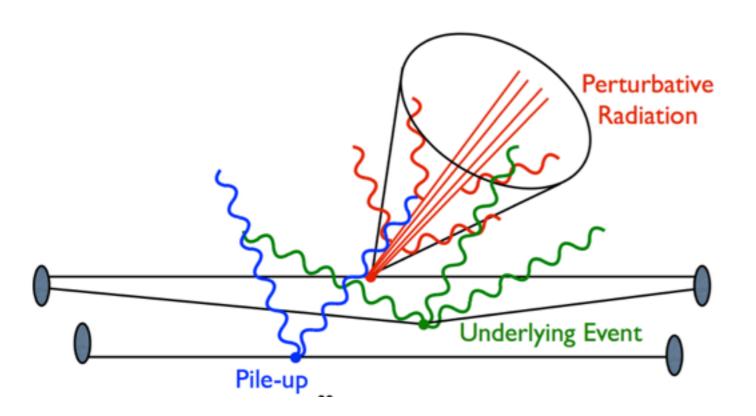
pruning: Ellis, Vermilion, Walsh 0912.0033

soft drop: Larkoski, Marzani,

Soyez, Thaler 1402.2657

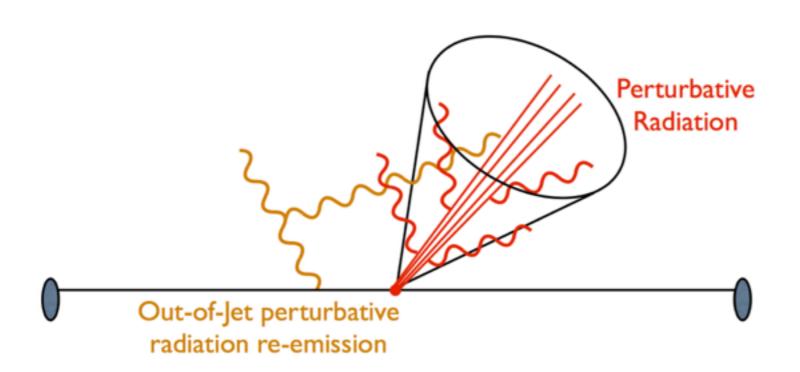
modified mass drop: Dasgupta,

Fregoso, Marzani, Salam 1307.0007



# Challenges in resuming jet substructure at NLL and beyond :

complication due to initial state radiation/ Non-global logs



Removing perturbative soft radiation:

final state wide-angle radiation

initial state radiation

non-global radiation

Reduces process dependence Eliminates non-global logarithms .....

Desired theoretical property for mMDT and soft drop:

No non-global logs at NLL, no dependence on jet radius.

[Dasgupta, Fregoso, Marzani, Salam 1307.0007] [Larkoski, Marzani, Soyez, Thaler 1402.2657]

#### Going beyond

- Can we prove the absence of non-global logs and jetradius dependence to all orders?
- What is the proper choice of grooming parameters/ reclustering algorithm that allow a clean theoretical description?

#### OUTLINE

- define soft-drop algorithm
- ➤ Factorization at e+e-
- ➤ Obtaining two-loop soft non-cusp anomalous dimension
- study of clustering effects

### THE SOFT-DROP GROOMER

#### **GROOMING PROCEDURE**

➤ 1. Take a jet with radius R~1, recluster with C/A.

$$d_{ij} = R_{ij} \equiv \sqrt{(y_i - y_j)^2 + (\phi_i - \phi_j)^2}$$

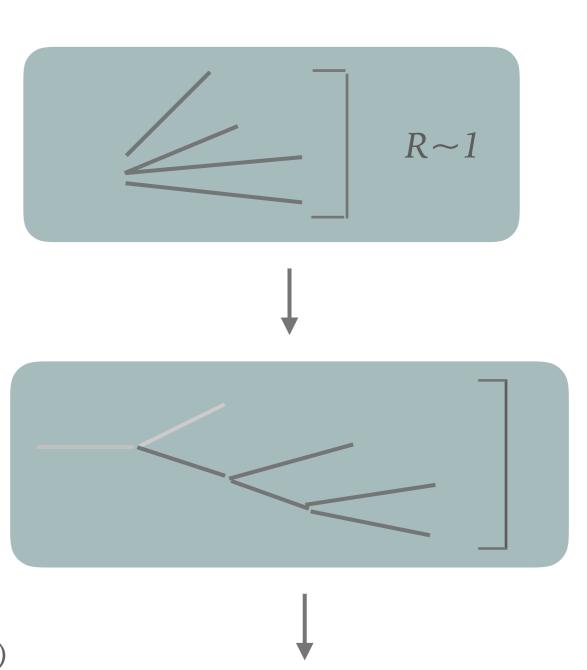
➤ 2 Undo the last step of reclustering, check if two branches (i,j) satisfy soft-drop condition. If not drop the softer brach.

$$\frac{\min[p_{Ti}, p_{Tj}]}{p_{Ti} + p_{Tj}} > z_{\text{cut}} \left(\frac{R_{ij}}{R}\right)^{\beta}$$

$$z_{\rm cut} \sim 0.1$$

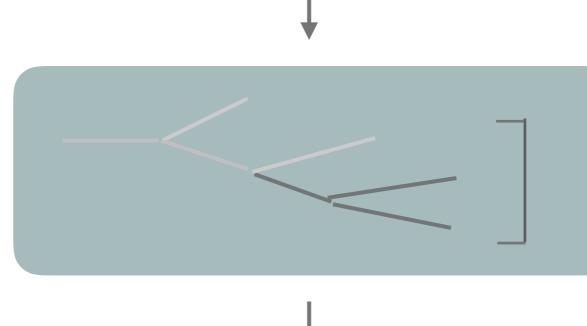
 $\beta = 0$  check only energy fraction (mMDT)

$$\beta = \infty$$
 no soft drop



#### **GROOMING PROCEDURE**

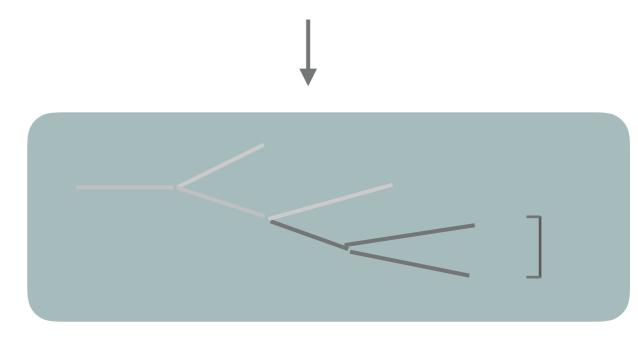
repeat step 2 on the harder branch, until the soft-drop condition is satisfied.



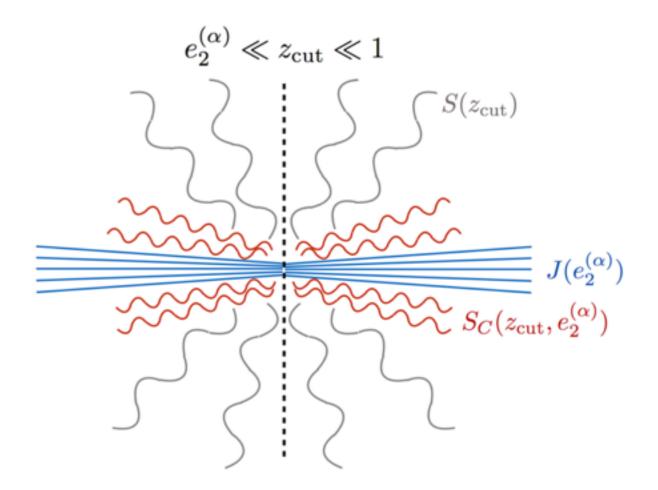
measure energy correlation function of the groomed jet

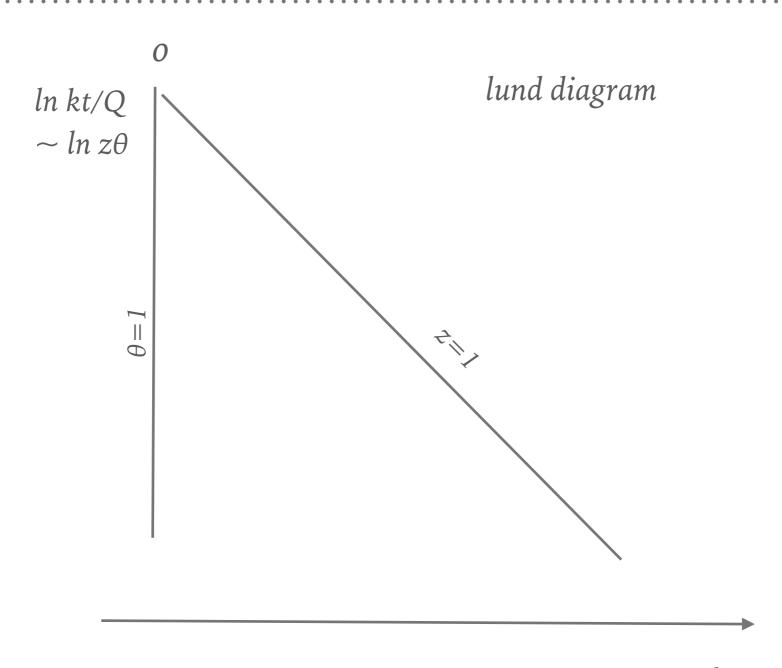
$$e_2^{(\alpha)}\Big|_{pp} = \frac{1}{p_{TJ}^2} \sum_{i < j \in J} p_{Ti} p_{Tj} R_{ij}^{\alpha}$$

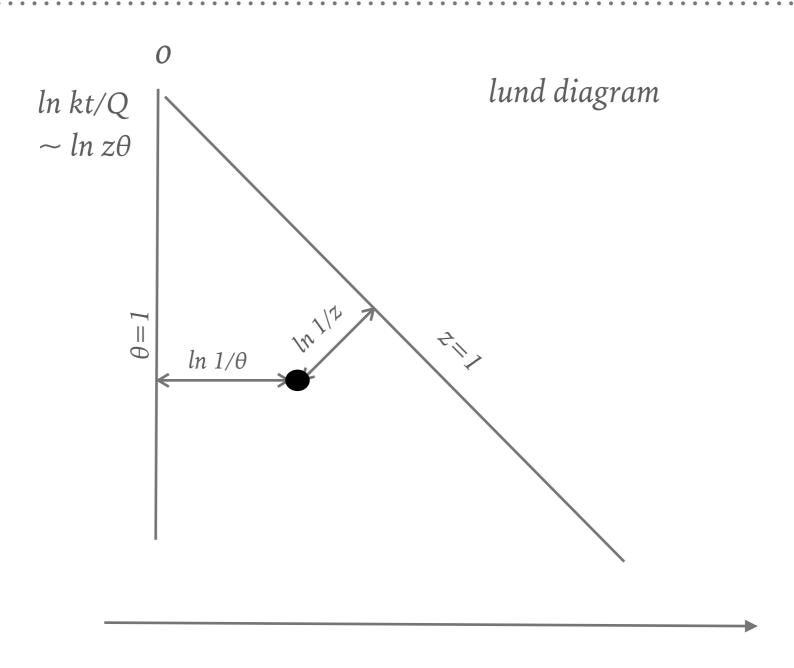
$$e_2^{(2)} = \frac{m_g^2}{E_{Jg}^2}$$

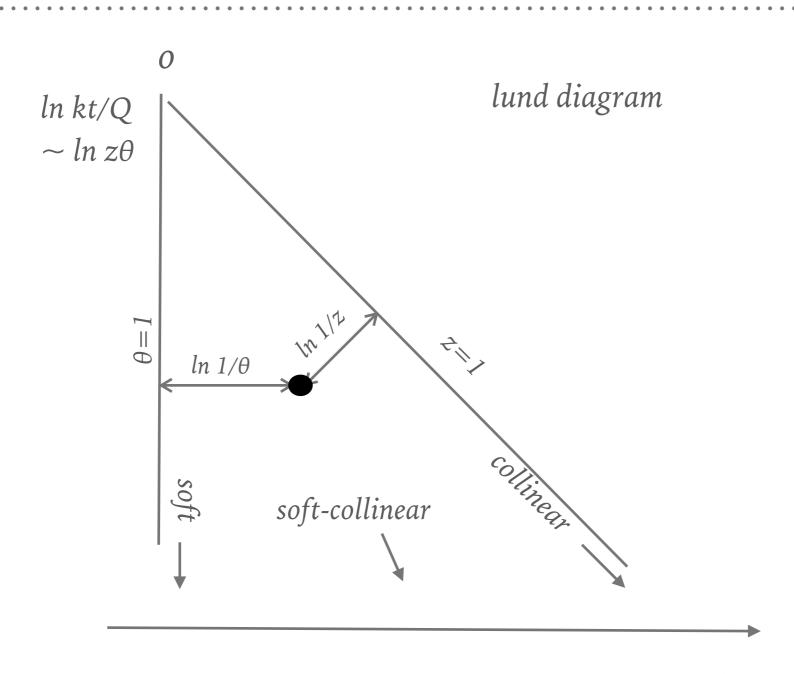


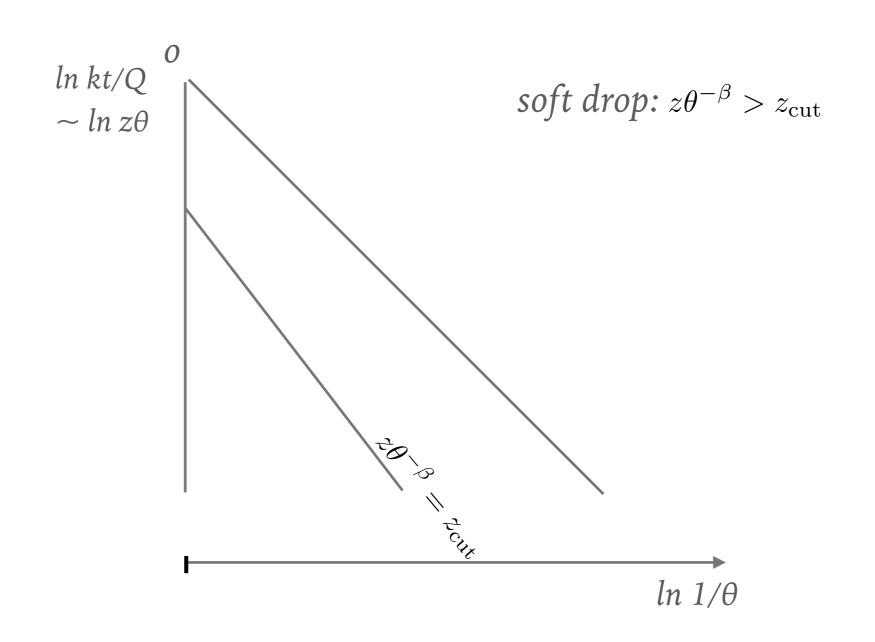
# FACTORIZATION AT E+E-—>DIJETS

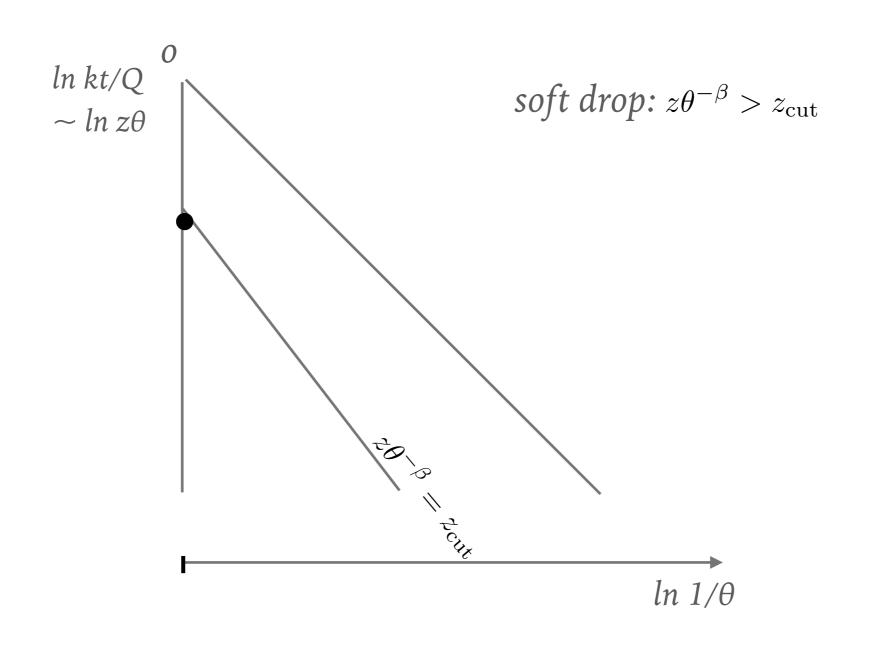


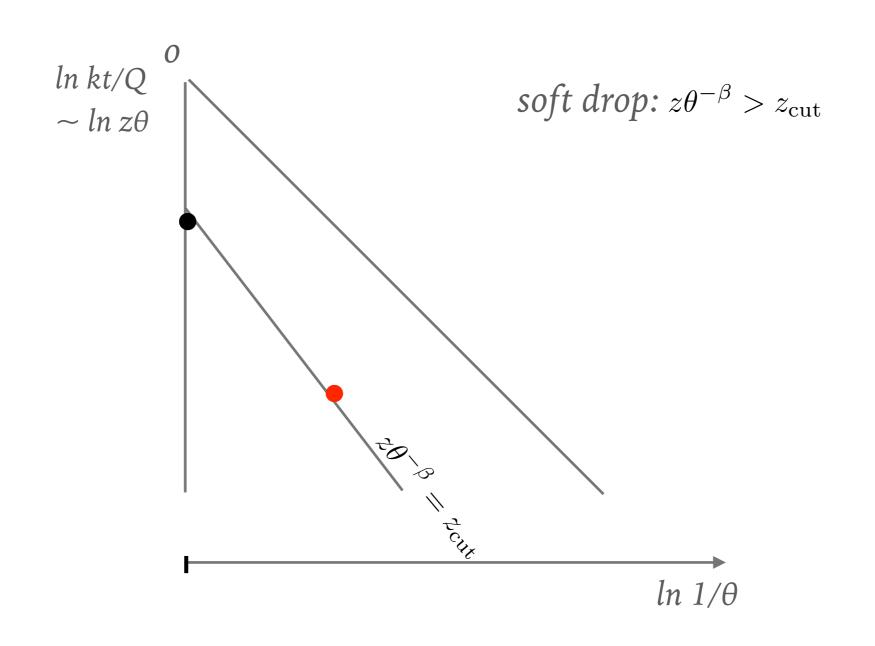


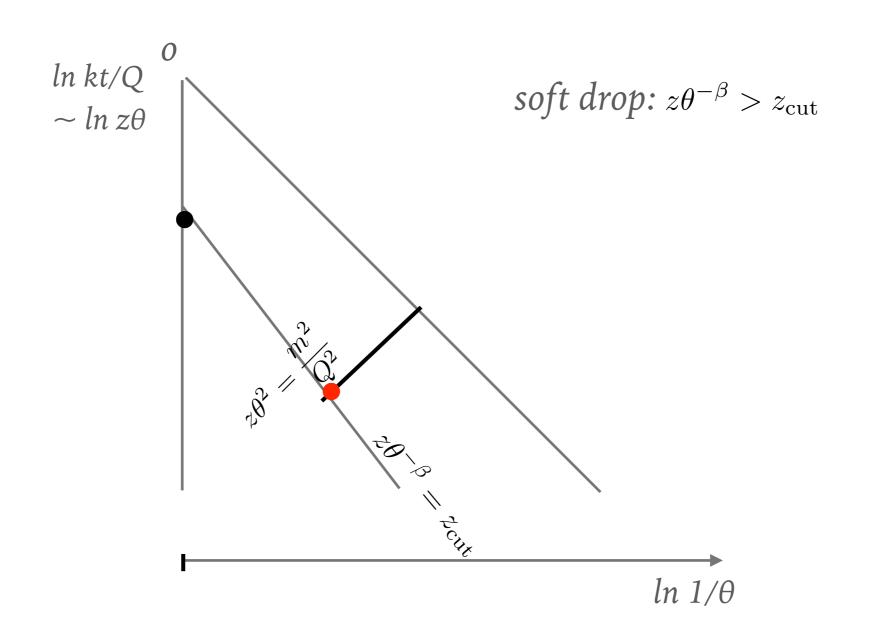


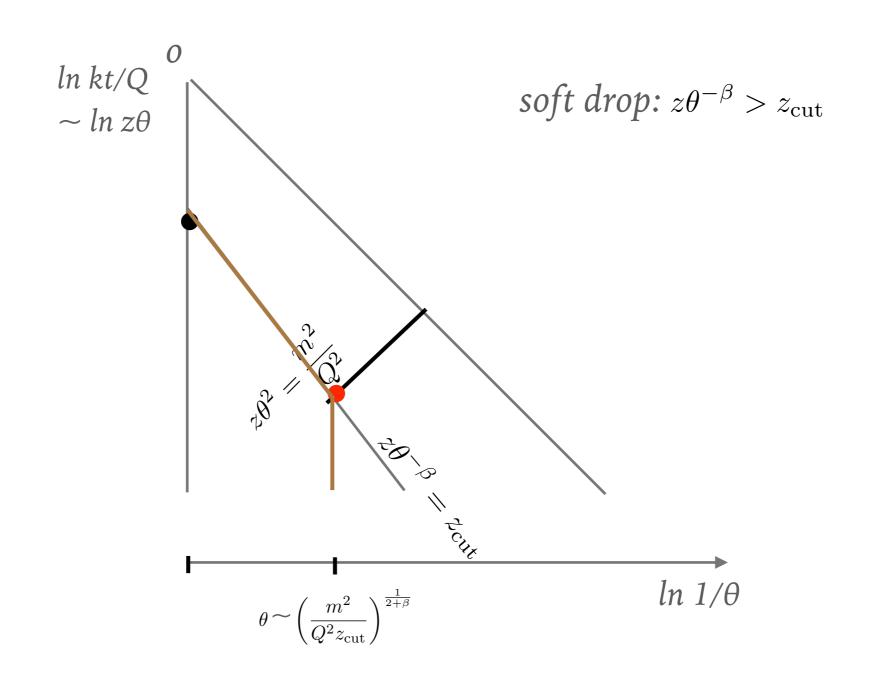


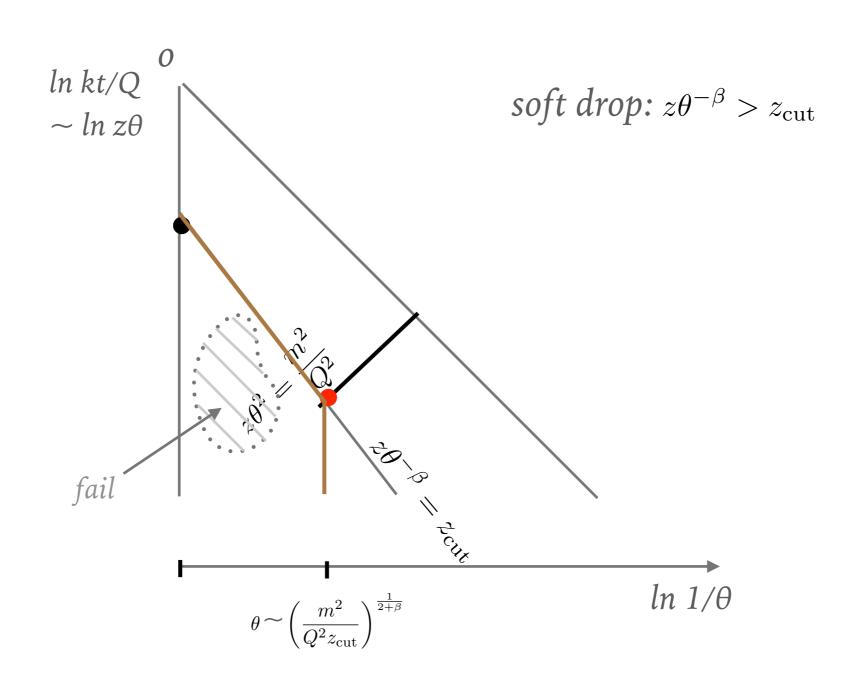


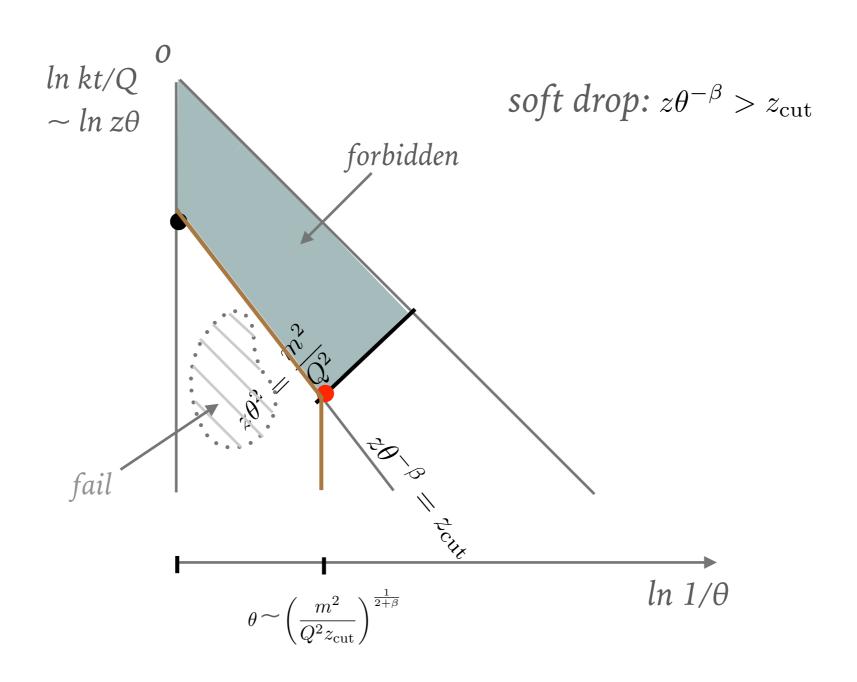


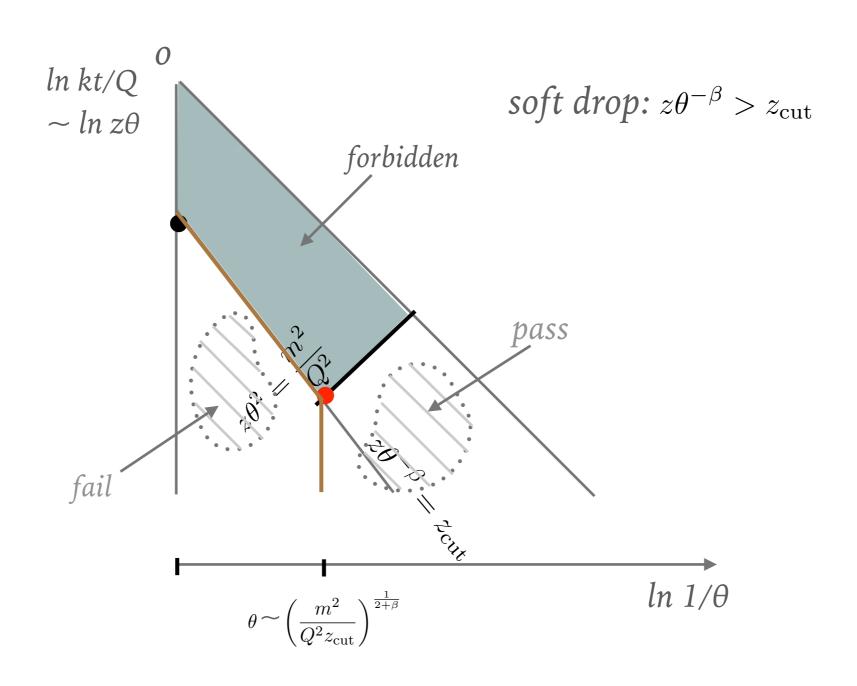


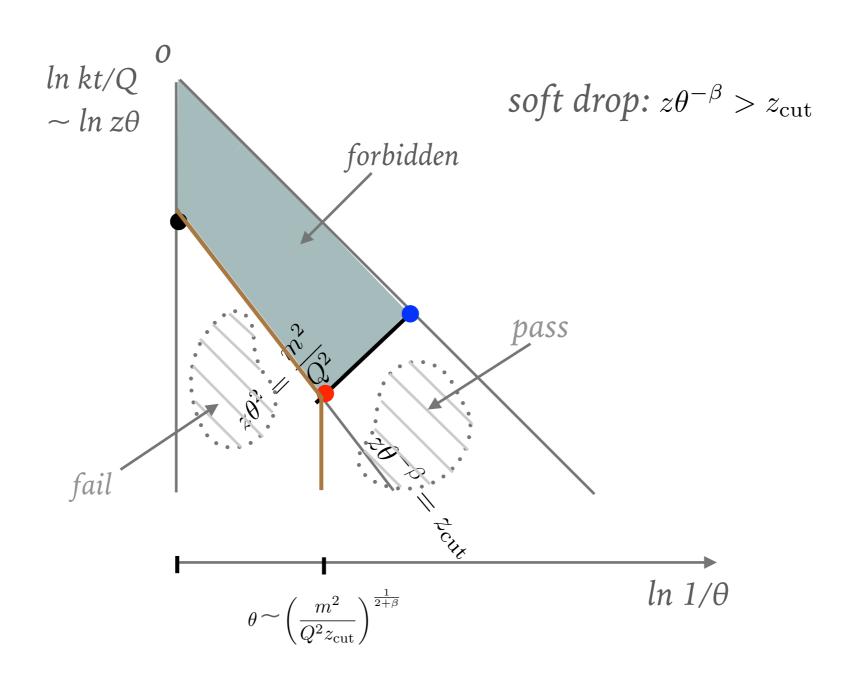












#### HIERARCHY OF SCALES

Focus on regime 
$$m^2/Q^2 \ll z_{\rm cut} \ll 1$$

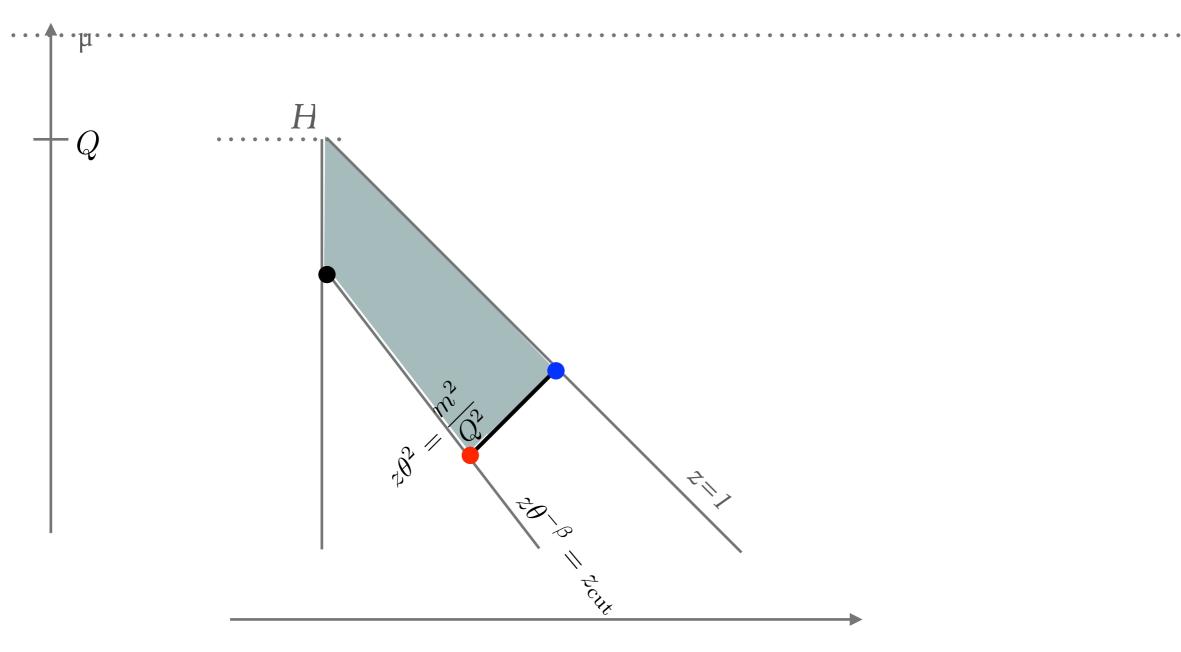
 $\rightarrow m^2/Q^2 \ll z_{\rm cut}$ 

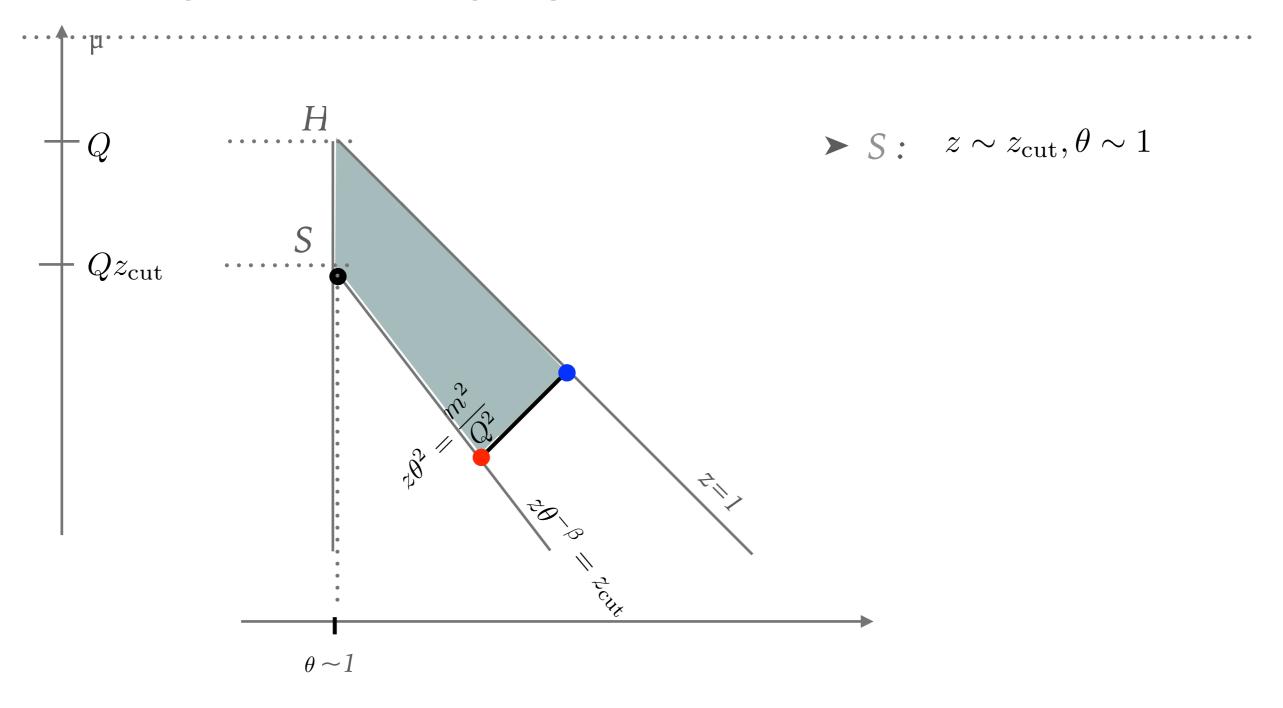
Wide angle emissions are not allowed to pass soft-drop

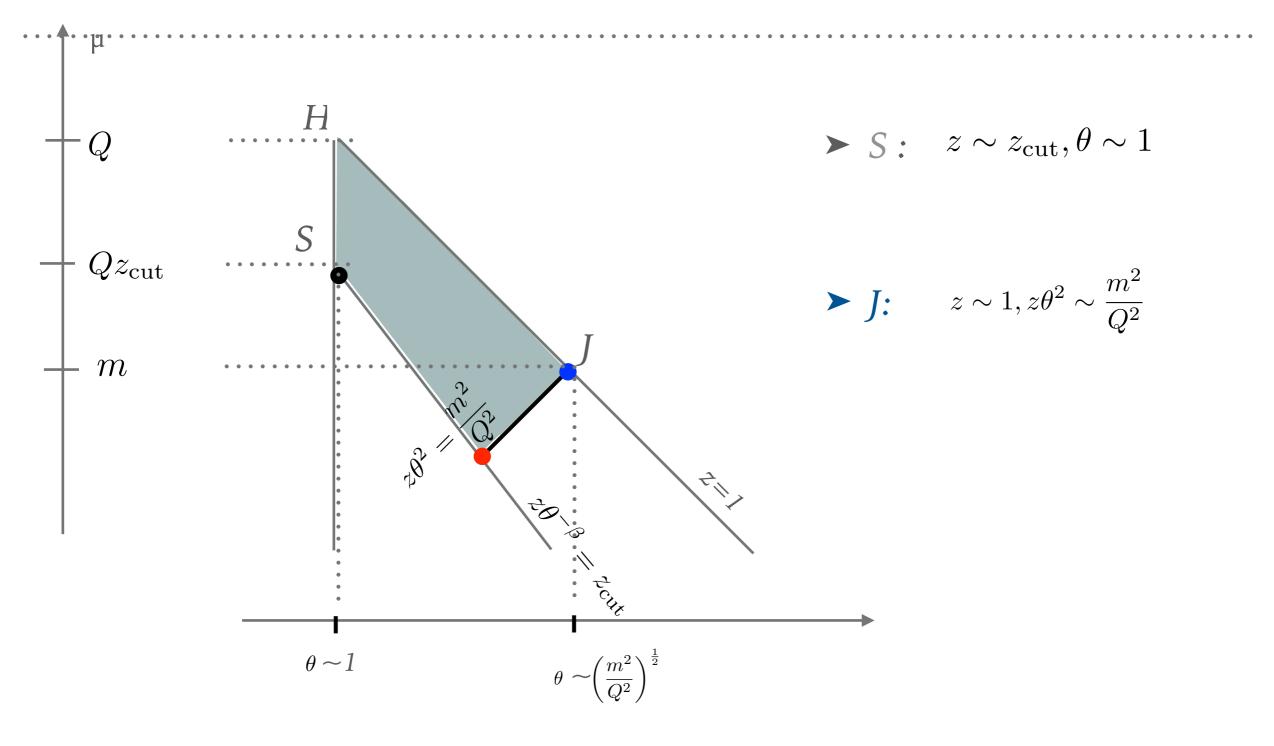
- After soft drop, all remaining particles in the jet must be collinear!
- $\rightarrow z_{\rm cut} \ll 1$

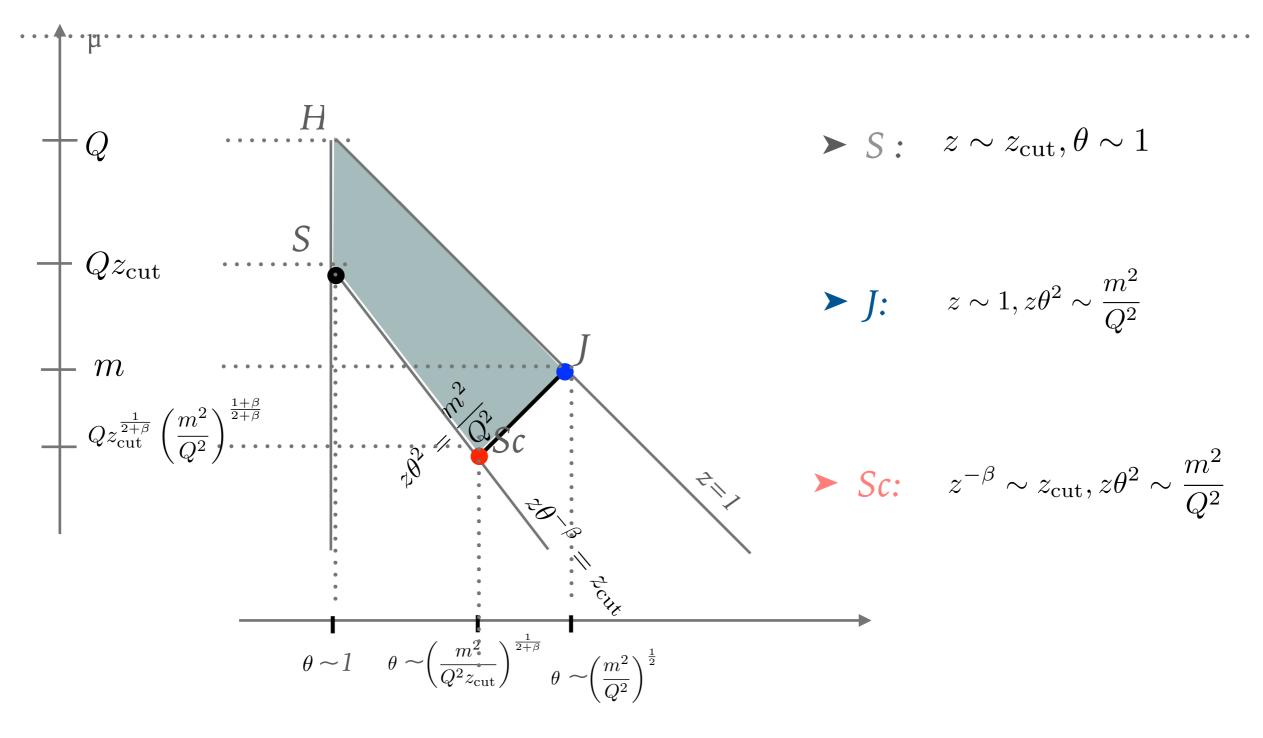
All hard-collinear particles survive soft-drop grooming.

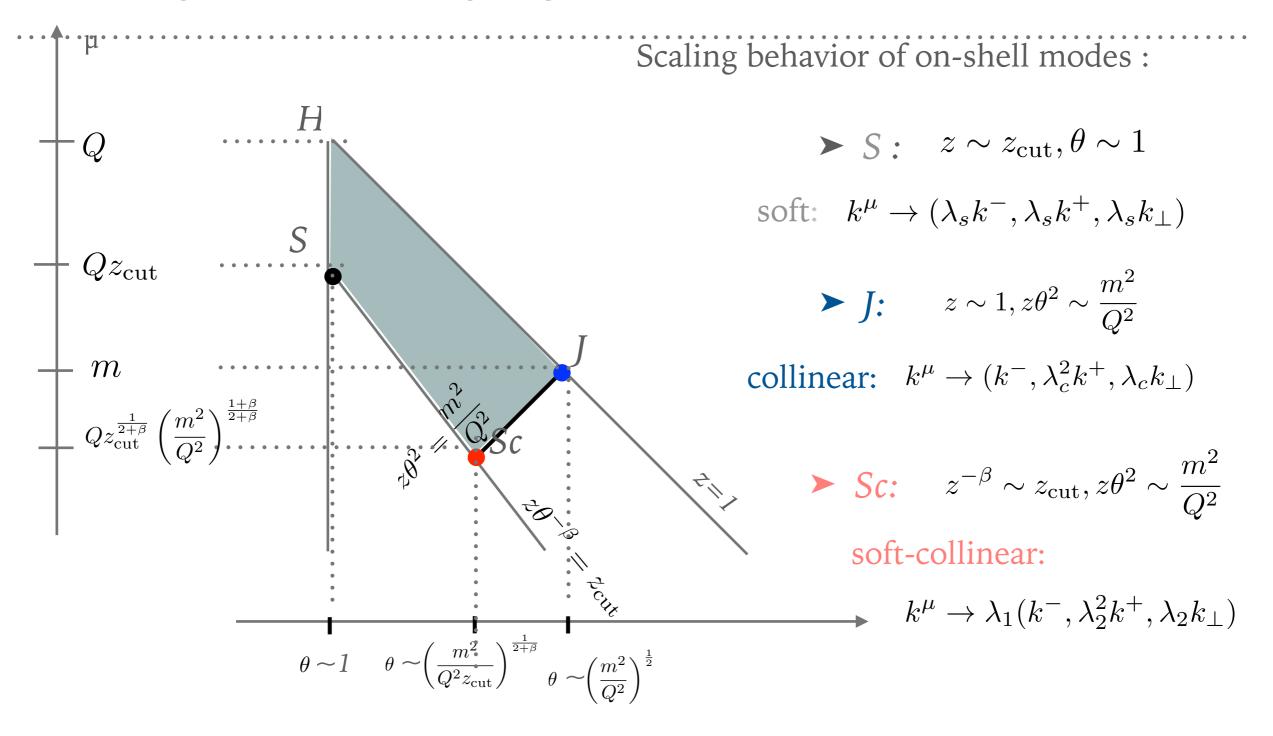
 effects of jet energy loss and change of jet flavor during grooming is power-suppressed by zcut. [Dasgupta, Fregoso, Marzani, Salam 1307.0007]



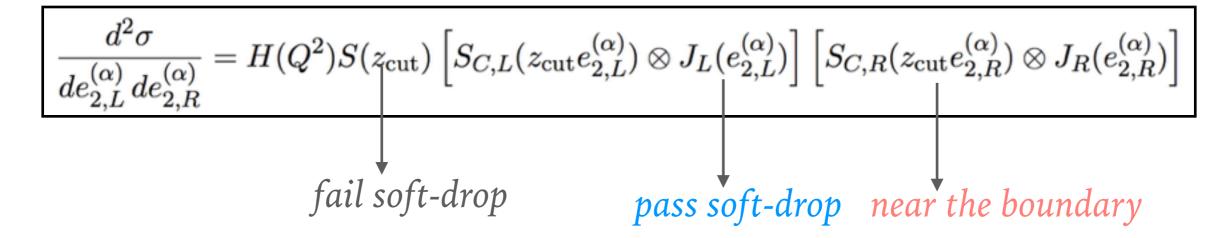








only collinear objects contribute to the groomed mass



- ➤ Consequences of factorization
- Collinear universality allows the predict the shape of the distribution at the LHC using results obtained at e+e-
- the shape is unaffected by non-global logarithms and is independent of the jet radius.
- (e+e-)single scale dependence: the absence of clustering logs/non-global logs

Check single scale dependence:

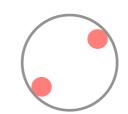
- The jet function is inclusive
- Can check either the soft function or softcollinear function depend on one single scale

 $\Theta_{SD}$ : C/A clustering + soft-drop veto

- ΘSD acts on soft/collinear/soft-collinear final state independently
- ➤ For C/A, clustering sequence is invariant under soft/collinear/soft-collinear scaling.

Check single scale dependence:

- The jet function is inclusive
- Can check either the soft function or softcollinear function depend on one single scale

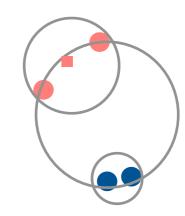


 $\Theta_{SD}$ : C/A clustering + soft-drop veto

- ΘSD acts on soft/collinear/soft-collinear final state independently
- ➤ For C/A, clustering sequence is invariant under soft/collinear/soft-collinear scaling.

#### Check single scale dependence:

- The jet function is inclusive
- Can check either the soft function or softcollinear function depend on one single scale



 $\Theta_{SD}$ : C/A clustering + soft-drop veto

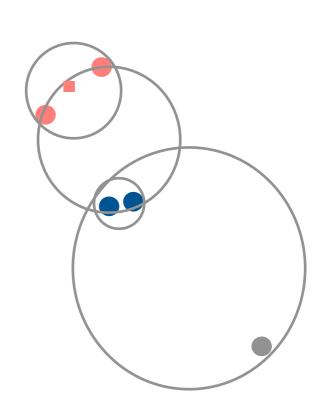
- ΘSD acts on soft/collinear/soft-collinear final state independently
- ➤ For C/A, clustering sequence is invariant under soft/collinear/soft-collinear scaling.

Check single scale dependence:

- The jet function is inclusive
- Can check either the soft function or softcollinear function depend on one single scale

 $\Theta_{SD}$ : C/A clustering + soft-drop veto

- ΘSD acts on soft/collinear/soft-collinear final state independently
- ➤ For C/A, clustering sequence is invariant under soft/collinear/soft-collinear scaling.

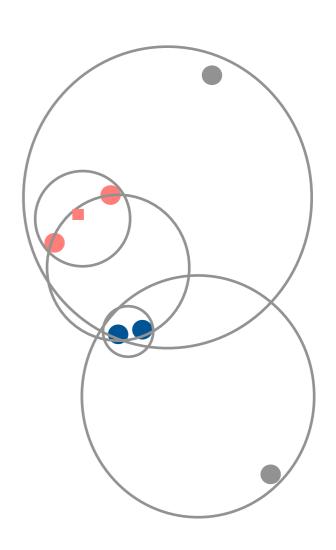


Check single scale dependence:

- The jet function is inclusive
- Can check either the soft function or softcollinear function depend on one single scale

 $\Theta_{SD}$ : C/A clustering + soft-drop veto

- ΘSD acts on soft/collinear/soft-collinear final state independently
- ➤ For C/A, clustering sequence is invariant under soft/collinear/soft-collinear scaling.

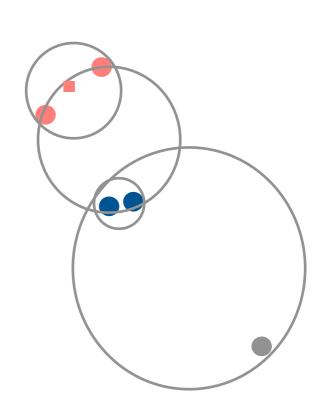


Check single scale dependence:

- The jet function is inclusive
- Can check either the soft function or softcollinear function depend on one single scale

 $\Theta_{SD}$ : C/A clustering + soft-drop veto

- ΘSD acts on soft/collinear/soft-collinear final state independently
- ➤ For C/A, clustering sequence is invariant under soft/collinear/soft-collinear scaling.

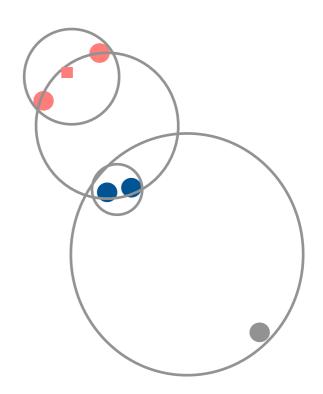


Check single scale dependence:

- The jet function is inclusive
- Can check either the soft function or softcollinear function depend on one single scale

 $\Theta_{SD}$ : C/A clustering + soft-drop veto

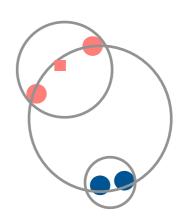
- ΘSD acts on soft/collinear/soft-collinear final state independently
- ➤ For C/A, clustering sequence is invariant under soft/collinear/soft-collinear scaling.



#### **CLUSTERING EFFECT**

#### Check single scale dependence:

- The jet function is inclusive
- Can check either the soft function or softcollinear function depend on one single scale



 $\Theta_{SD}$ : C/A clustering + soft-drop veto

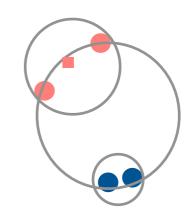
#### Strict factorization holds

- ΘSD acts on soft/collinear/soft-collinear final state independently
- ➤ For C/A, clustering sequence is invariant under soft/collinear/soft-collinear scaling.

#### **CLUSTERING EFFECT**

Check single scale dependence:

- The jet function is inclusive
- Can check either the soft function or softcollinear function depend on one single scale



 $\Theta_{SD}$ : C/A clustering + soft-drop veto

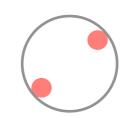
Strict factorization holds

- ΘSD acts on soft/collinear/soft-collinear final state independently
- ➤ For C/A, clustering sequence is invariant under soft/collinear/soft-collinear scaling.

#### **CLUSTERING EFFECT**

Check single scale dependence:

- The jet function is inclusive
- Can check either the soft function or softcollinear function depend on one single scale



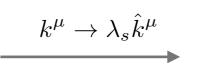
 $\Theta_{SD}$ : C/A clustering + soft-drop veto

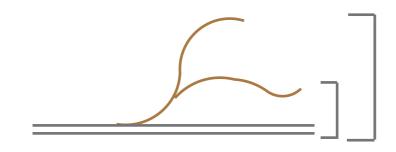


#### Strict factorization holds

- ΘSD acts on soft/collinear/soft-collinear final state independently
- ➤ For C/A, clustering sequence is invariant under soft/collinear/soft-collinear scaling.

$$S(z_{\rm cut}) = \sum_{n} \mu^{2n\epsilon} \int d\Pi_n |\mathcal{M}_n|^2 \Theta_{\rm SD}$$





$$\Theta(\theta_{ij} - \min[\theta_i, \theta_j])$$
 -----

$$\Theta(\hat{\theta}_{ij} - \min[\hat{\theta}_i, \hat{\theta}_j])$$

soft drop

$$\Theta(z_{\rm cut}\theta^{\beta} - \sum_{i} z_{i})$$

$$\Theta(\theta^{\beta} - \sum_{i} z_{i})$$

SCET matrix element

$$d\Pi_n |\mathcal{M}_n|^2$$

$$z_{\mathrm{cut}}^{-2n\epsilon}d\Pi_n|\mathcal{M}_n|^2$$

$$S(z_{\text{cut}}) = \sum_{n} \mu^{2n\epsilon} (z_{\text{cut}})^{-2n\epsilon} \int d\Pi_{n} |\mathcal{M}_{n}|^{2} \Theta_{\text{SD}}^{z_{\text{cut}}=1}$$
 depends on a single soft scale Qzcut.

By similar argument,

$$S_C(z_{\rm cut}m^2) = \sum_n \mu^{2n\epsilon} \left( z_{\rm cut}^{\frac{1}{2+\beta}} (m^2)^{\frac{1+\beta}{2+\beta}} \right)^{-2n\epsilon} \frac{Q^2}{m^2} \int d\Pi_n |\mathcal{M}_n|^2 \Theta_{SD}^{z_{\rm cut}=1} \delta_{\frac{m^2}{Q^2}=1}$$

#### **ACHIEVING NNLL ACCURACY**

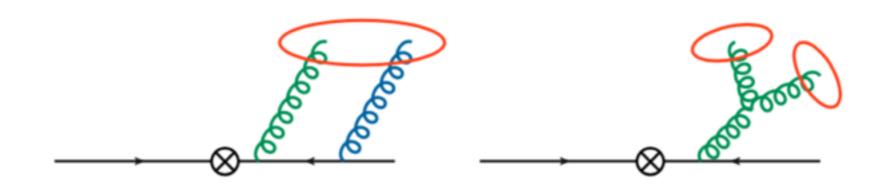
missing ingredient: two-loop non cusp anomalous dimension of the soft-collinear function.

$$\frac{d^2\sigma}{de_{2,L}^{(\alpha)}de_{2,R}^{(\alpha)}} = H(Q^2)S(z_{\text{cut}}) \left[ S_{C,L}(z_{\text{cut}}e_{2,L}^{(\alpha)}) \otimes J_L(e_{2,L}^{(\alpha)}) \right] \left[ S_{C,R}(z_{\text{cut}}e_{2,R}^{(\alpha)}) \otimes J_R(e_{2,R}^{(\alpha)}) \right]$$

$$0 = \gamma_H + \gamma_S + 2\gamma_J + 2\gamma_{S_C}$$

- ➤ calculate the soft anomalous dimension from two-loop hemisphere soft function at e+e-:
- two wilson-line fixed angle matrix element
- simple phase-space constraint
- for  $\beta = 0$ , related to calculations from literature [Manteuffel, Schabinger, Zhu, 1309.3560];
- for  $\beta = 1$ , we extracted the anomalous dimension from EVENT2

# **GROOMING CALCULATION AT NNLO**



#### TWO-LOOP HEMISPHERE SOFT FUNCTION

can be related to the two-loop soft-function with a global energy veto

[Manteuffel, Schabinger, Zhu, 1309.3560]

$$S_{\text{veto}} = \int d\Pi_2 |\mathcal{M}(k_1, k_2)|^2 \Theta(\frac{Q}{2}z \text{cut} - k_1^0 - k_2^0)$$

Two-loop non-cusp anomalous dimension

$$\gamma_{\rm SD} = \gamma_{\rm veto} + \gamma_{\rm Clust}^{\rm alg.}$$

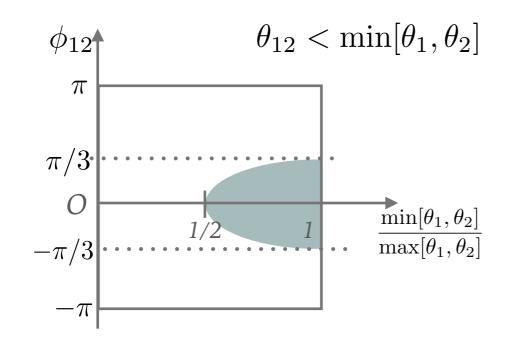
- ➤ Abelian contribution: purely comes from clustering effect that violates abelian exponentiation
- ➤ Non-abelian contribution:

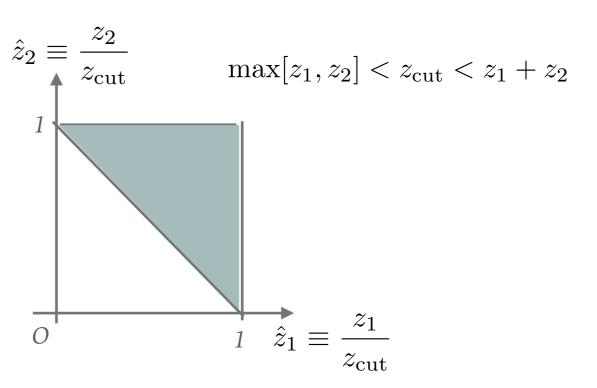
only need to consider the difference between soft drop and energy veto phase-space constraints, which is a highly restricted phase-space region.

## INDEPENDENT EMISSIONS

 $S(z_{
m cut})|_{
m A,lpha_s^2}$  +  $\left[S^{1m loop}(z_{
m cut})\right]^2$  clustering correction

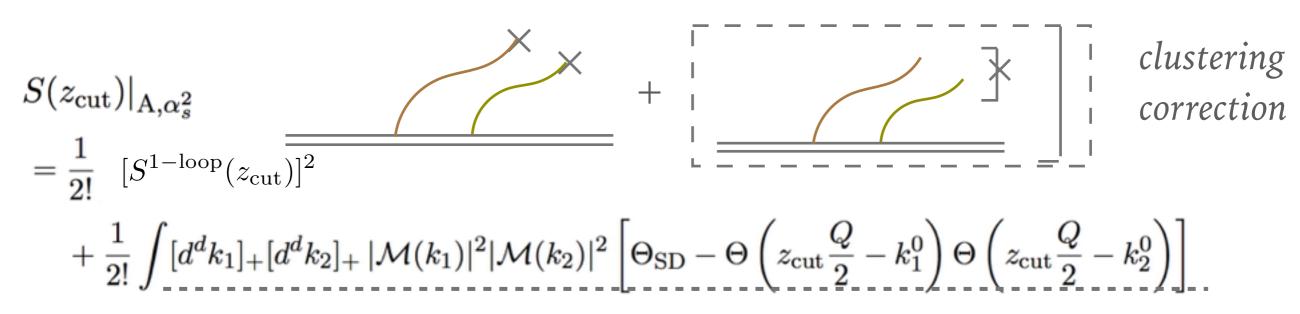
$$+\frac{1}{2!} \int [d^d k_1]_+ [d^d k_2]_+ |\mathcal{M}(k_1)|^2 |\mathcal{M}(k_2)|^2 \left[\Theta_{SD} - \Theta\left(z_{\text{cut}} \frac{Q}{2} - k_1^0\right) \Theta\left(z_{\text{cut}} \frac{Q}{2} - k_2^0\right)\right]$$

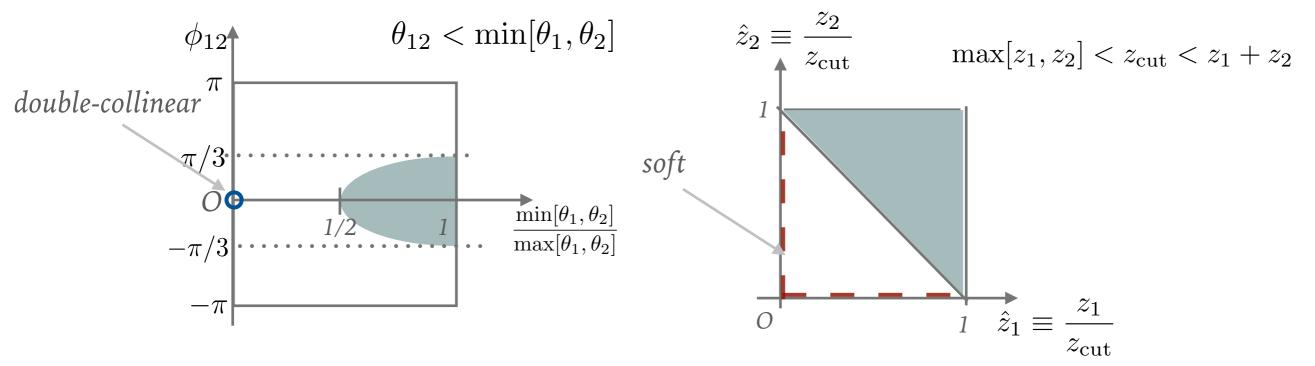




$$\gamma_{\text{C/A}}^{\text{A},\alpha_s^2} = \left(\frac{\alpha_s}{4\pi}\right)^2 34.01 \, C_F^2$$

### INDEPENDENT EMISSIONS

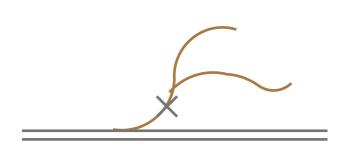




$$\gamma_{\text{C/A}}^{\text{A},\alpha_s^2} = \left(\frac{\alpha_s}{4\pi}\right)^2 34.01 \, C_F^2$$

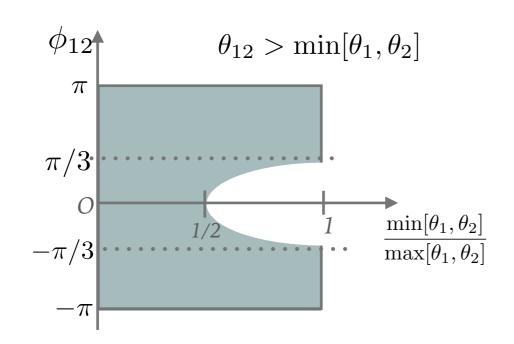
### **CORRELATED EMISSIONS**





clustering correction

$$S(z_{\text{cut}})|_{\text{n-A},\alpha_s^2} = S_{\text{veto}}|_{\text{n-A},\alpha_s^2} + \int [d^d k_1]_+ [d^d k_2]_+ |\mathcal{M}_{\text{n-A}}(k_1,k_2)|^2 [\Theta_{\text{SD}} - \Theta_{\text{veto}}]$$



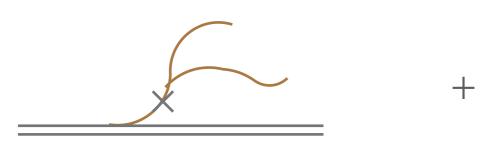
$$\hat{z}_2 \equiv \frac{z_2}{z_{\text{cut}}} \qquad \max[z_1, z_2] < z_{\text{cut}} < z_1 + z_2$$

$$1 \qquad 1 \qquad \hat{z}_1 \equiv \frac{z_1}{z_{\text{cut}}}$$

$$\gamma_{\text{C/A}}^{\text{n-A},\alpha_s^2} = \left(\frac{\alpha_s}{4\pi}\right)^2 C_F \left[-9.31C_A - 14.04n_f T_R\right]$$

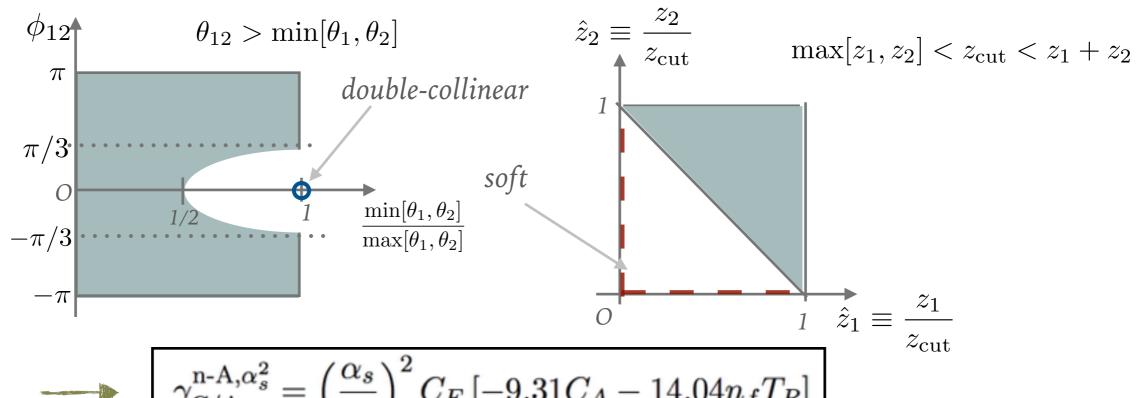
### **CORRELATED EMISSIONS**





clustering correction

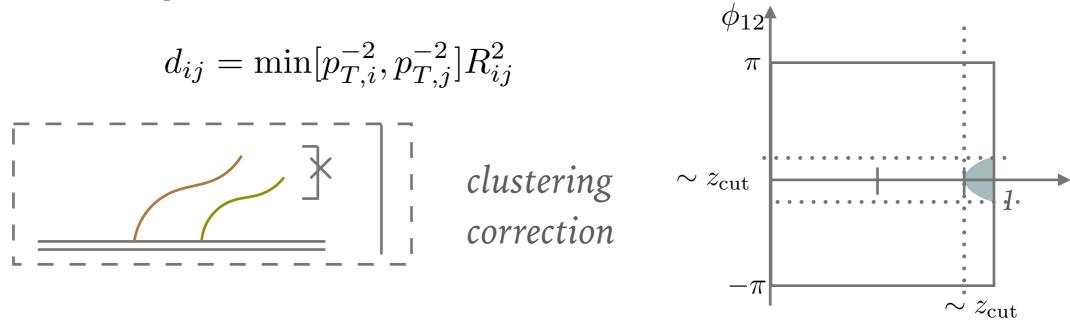
$$S(z_{\text{cut}})|_{\text{n-A},\alpha_s^2} = S_{\text{veto}}|_{\text{n-A},\alpha_s^2} + \int [d^d k_1]_+ [d^d k_2]_+ |\mathcal{M}_{\text{n-A}}(k_1,k_2)|^2 [\Theta_{\text{SD}} - \Theta_{\text{veto}}]$$



$$\gamma_{\text{C/A}}^{\text{n-A},\alpha_s^2} = \left(\frac{\alpha_s}{4\pi}\right)^2 C_F \left[-9.31C_A - 14.04n_f T_R\right]$$

#### RECLUSTERING WITH ANTI-KT ALGORITHM

Clustering metric



cluster soft emissions i,j before clustering them with the jet only if

$$\min[\hat{z}_i^{-2}, \hat{z}_j^{-2}]\theta_{ij}^2 < z_{\text{cut}}^2 \min[\theta_i^2, \theta_j^2]$$

The probability of clustering independent emissions together is suppressed by zcut ^ 2.

➤ No abelian clustering correction at leading power.

The non-abelian clustering logs can be obtained by computing the correction to  $S_{\text{veto}}$ 

$$\gamma_{ak_T} = -8\left(\frac{\alpha_s}{4\pi}\right)^2 C_F \left\{ \left[ \left( \frac{131}{9} - \frac{4}{3}\pi^2 - \frac{44}{3}\log 2 \right) C_A + \left( -\frac{46}{9} + \frac{16}{3}\log 2 \right) n_f T_R \right] \log z_{\text{cut}} \right.$$

$$\left. + \left( -\frac{269}{6} + \frac{7}{2}\zeta_3 + \frac{274}{9}\log 2 + \frac{11\pi^2}{9} + \frac{44}{3}\log^2 2 \right) C_A$$

$$\left. + \left( \frac{53}{3} - \frac{4\pi^2}{9} - \frac{116}{9}\log 2 - \frac{16}{3}\log^2 2 \right) n_f T_R \right\}$$

identical to the leading clustering log in jet-

[Tackmann, Walsh, Zuberi 1206.4312][Banfi, Salam, Zanderighi 1203.5773]

clustering log in jet-
veto calculation 
$$C_2(R) = 2C_A \left[ \left( 1 - \frac{8\pi^2}{3} \right) C_A + \left( \frac{23}{3} - 8 \ln 2 \right) \beta_0 \right] \ln R^2$$

[Tackmann, Walsh, Zuberi  $+ 15.62C_A^2 - 9.17C_A\beta_0 + C_2^{Rsub}(R)$ 

### RELATING TO JET VETO CALCULATION

➤ Clustering metric at the LHC with anti-kT algorithm: cluster two soft emissions into one jet if  $d_{ij}^{\text{eff}} < 1$ 

$$d_{ij}^{\text{eff},pp} \equiv \min[\hat{z}_i^{-2}, \hat{z}_j^{-2}] \frac{R_{ij}^2}{R^2}$$

boost-invariant geometrical separation phase space

Effective clustering metric at e+e-: combine two soft emissions (i,j) before combining them with hard jet core if  $d_{ij}^{\text{eff}} < 1$ 

$$d_{ij}^{\text{eff},e^+e^-} \equiv \min[\hat{z}_i^{-2},\hat{z}_j^{-2}] \frac{\theta_{ij}^2}{z_{\text{cut}}^2 \min[\theta_i^2,\theta_j^2]}$$

boost-invariant angle along the jet axis

➤ In large rapidity region (where log R arises), set R=zcut,

$$d_{ij}^{\text{eff},pp} \xrightarrow{i \parallel j} d_{ij}^{\text{eff},e^+e^-}$$

#### **SUMMARY**

- > grooming parameter and measuring regime:  $z_{\rm cut} \sim 0.1~m^2/Q^2 \ll z_{\rm cut}$
- preferred reclustering algorithm: C/A

With the above choices, we obtained:

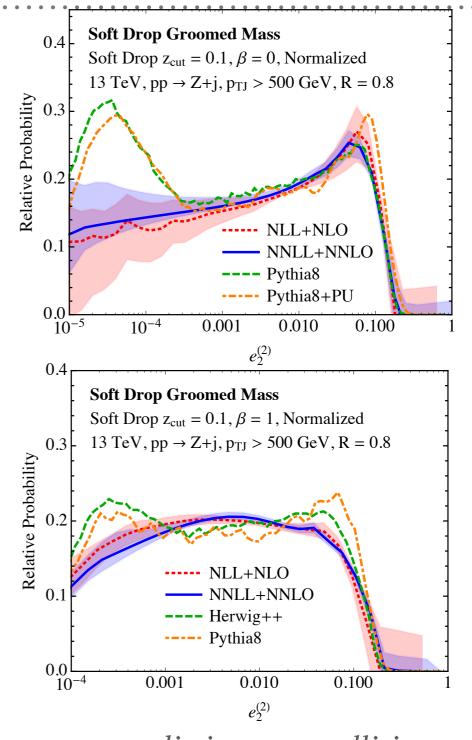
- All-order factorization at leading power of pile-up insensitive, process independent jet observable.
- First calculation done at NNLL accuracy with no non-global logs of jet substructure observable

#### **SUMMARY**

What can be done in the future:

- sub-leading power factorization theorem needed to predict Zcut power-corrections
- Soft-drop observables that extract more information on jet substructure:
   D2 etc;
- precision grooming calculation:

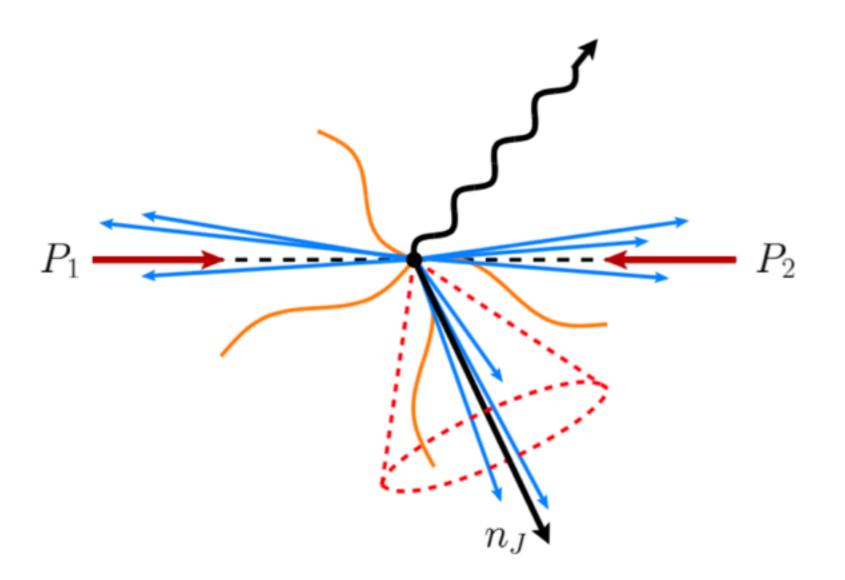
going to N3LO only requires two-Wilson line soft function calculation



prediction at pp collision

(to be continued...)

# THANKS!



### COMPLICATIONS AT HADRON COLLIDERS

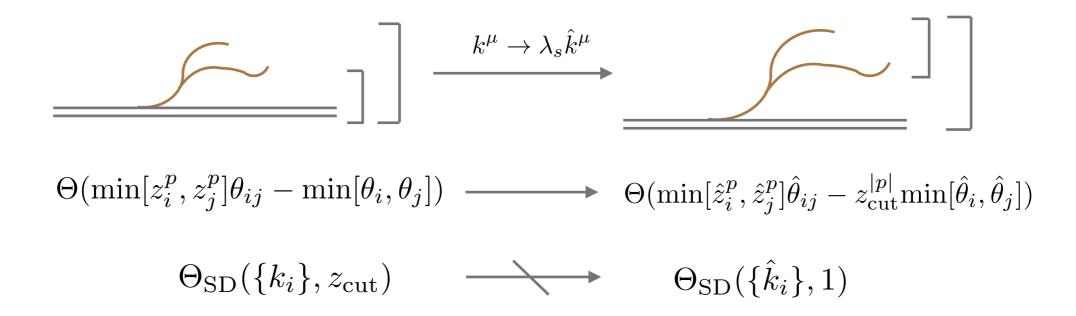
Factorization formula (in general, at leading z\_cut)

$$\frac{d\sigma_{\text{resum}}}{de_2^{(\alpha)}} = \sum_{k=q,\bar{q},g} D_k(p_T^{\min}, \eta_{\max}, z_{\text{cut}}, R) S_{C,k}(z_{\text{cut}} e_2^{(\alpha)}) \otimes J_k(e_2^{(\alpha)})$$

## **CLUSTERING LOGS (ANTI-KT ALGORITHM)**

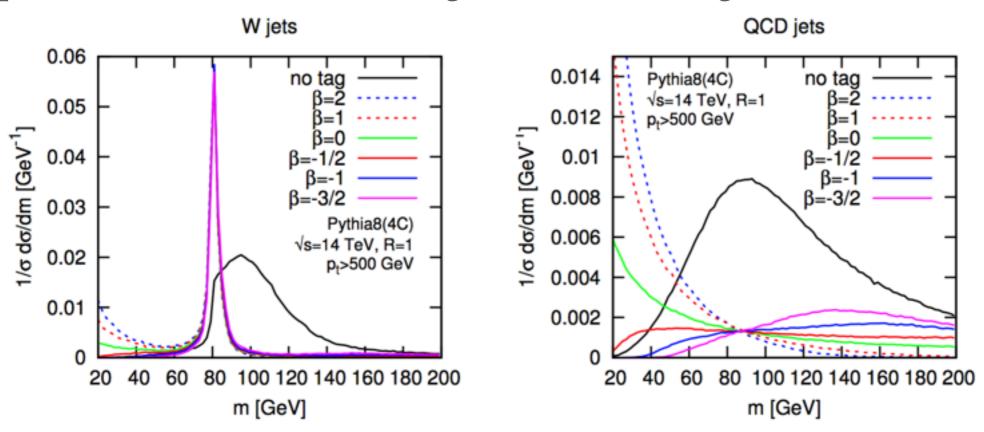
- study the clustering effect by looking at a different algorithm: anti-kT.
- Two soft emissions are clustered together before clustering with hard-collinear emission if  $\min[z_i^{-1}, z_j^{-1}]\theta_{ij} < \min[\theta_i, \theta_j]$

This inequality is not preserved after rescaling the soft momenta.

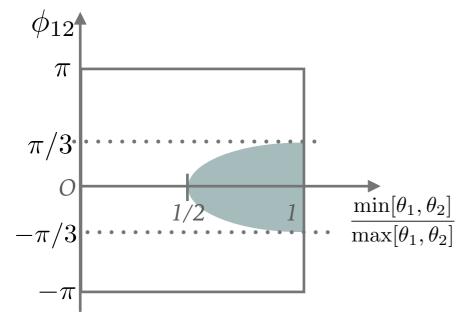


➤ If two independent emissions are clustered into the same jet then the rate is suppressed by the area of the jet which is O(zcut). At leading power, this implies the absence of abelian clustering effect.

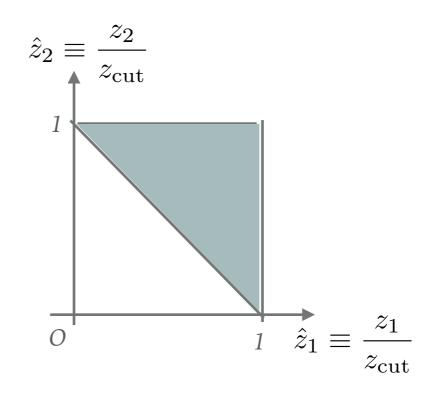
> Experimental control of signal over background



Performance of soft drop as a boosted W tagger: mass distribution of signal (left) and background (right) jets before and after soft drop [Larkoski, Marzani, Soyez, Thaler 1402.2657]



$$\Theta_{
m veto} = \Theta \left( \Lambda - k_1^0 - k_2^0 \right)$$



$$\Theta_{\text{SD}} - \Theta_{\text{veto}} = \left\{ \Theta(\eta_1 \eta_2) \left[ 1 - \Theta(\theta_{1J} - \theta_{12}) \Theta(\theta_{2J} - \theta_{12}) \right] + \Theta(-\eta_1 \eta_2) \right\} 
\times \Theta\left( z_{\text{cut}} \frac{Q}{2} - k_1^0 \right) \Theta\left( z_{\text{cut}} \frac{Q}{2} - k_2^0 \right) \Theta\left( k_1^0 + k_2^0 - z_{\text{cut}} \frac{Q}{2} \right) 
= -\Theta(\eta_1 \eta_2) \Theta(\theta_{1J} - \theta_{12}) \Theta(\theta_{2J} - \theta_{12}) 
\times \Theta\left( z_{\text{cut}} \frac{Q}{2} - k_1^0 \right) \Theta\left( z_{\text{cut}} \frac{Q}{2} - k_2^0 \right) \Theta\left( k_1^0 + k_2^0 - z_{\text{cut}} \frac{Q}{2} \right)$$