Factorization of Jet Substructure Observables

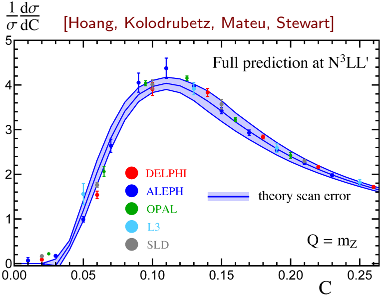
Ian Moult

Massachusetts Institute of Technology

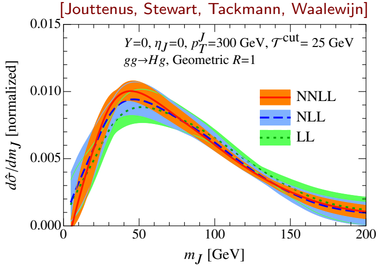
With Andrew Larkoski and Duff Neill

Motivation

- Inclusive properties of events (thrust, C-parameter,...) and jets (mass, angularities,...) "well" understood: all orders factorization theorems, resummation to high orders, treatment of NP physics, ...
- Allows precision comparison to data and Monte Carlo programs.



C-Parameter in e^+e^-



Jet Mass in $gg \to Hg$

Jet Substructure Revolution

- Higher energies and better detectors

  $\implies$ more detailed questions are being asked, and are important for optimizing reach of LHC.

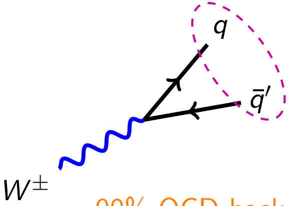
- Plethora of substructure observables proposed and measured

  $\tau_{2,1}^{(\beta)}$, $\tau_{3,2}^{(\beta)}$, $C_2^{(\beta)}$, Γ_{Qjet}, pruned/trimmed/... masses, ...

  [For a review see: 1311.2708]

- Exciting, BUT
  - Studied (almost*) purely with Monte Carlo.
  - Limited input on structure of variables from factorization considerations.

*See: [Feige, Schwartz, Stewart, Thaler]
[Dasgupta, Fregoso, Marzani, Salam]
[Dasgupta, Powling, Siodmok]



$\sim$ 90% QCD background rejection at 50% acceptance.

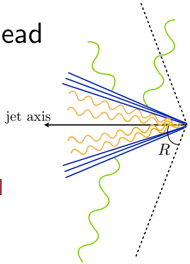
[CMS 1410.4227]

Motivation

- Lots to be learned in analytic Jet Substructure!

- Study of factorization properties leads to new, improved observables/ algorithms. e.g. D_2, D_3
  [mMDT: Dasgupta, Fregoso, Marzani, Salam]
  [Soft Drop: Larkoski, Marzani, Soyez, Thaler]
  [D2,D3: Larkoski, IM, Neill]

- Analytic calculations of substructure observables can lead to improved understanding of QCD factorization:
  - Additional hierarchical scales. [Bauer, Tackmann, Walsh, Zuberi]
  - Non-Global logarithms. [Dasgupta, Salam]
  - Sudakov safety of ratio observables. [Larkoski, Marzani, Thaler]

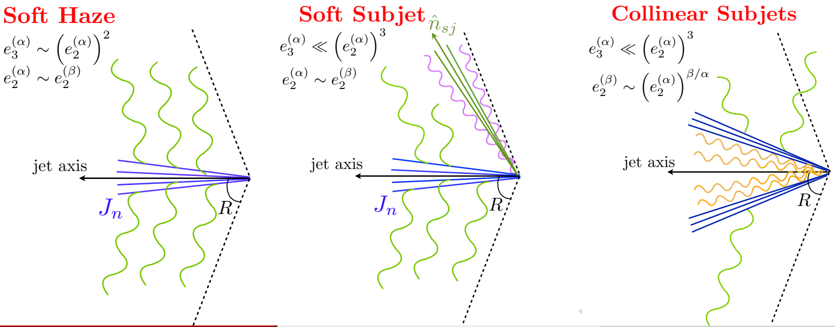


- Understanding of more detailed correlations important for improving Monte Carlo description of QCD shower.

Outline

- Use 2-prong discriminants as the simplest example of a substructure observable in which subjets are resolved.
- Describe how power counting in the EFT naturally identifies optimal discriminating observables. [Larkoski, IM, Neill 1409.6298, 1411.0665]
- Factorization of the D_2 observable: description of relevant phase space. [Larkoski, IM, Neill 1501.04596, 1504.soon]

[Bauer, Tackmann, Walsh, Zuberi]

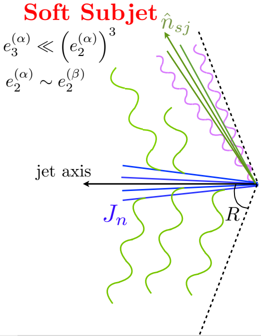


Outline

- Use 2-prong discriminants as the simplest example of a substructure observable in which subjets are resolved.
- Describe how power counting in the EFT naturally identifies optimal discriminating observables. [Larkoski, IM, Neill 1409.6298, 1411.0665]
- Factorization of the D_2 observable: description of relevant phase space. [Larkoski, IM, Neill 1501.04596, 1504.soon]

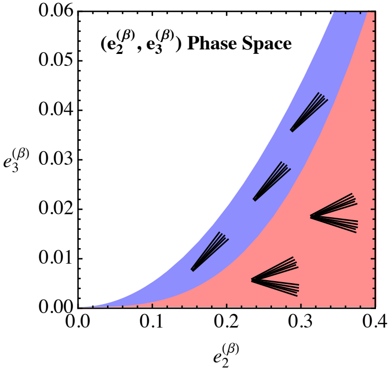
Duff Neill's Talk:

- Use same observables/phase space to discuss the resummation of NGLs.



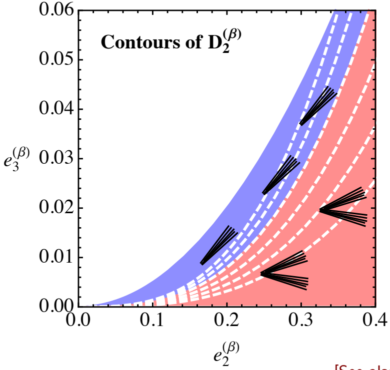
What is a 2-Prong Discriminant?

- Measure some set of observables on a jet, call them $e_2^{(\beta)}$, $e_3^{(\beta)}$.
- One and two prong jets live in distinct regions of the phase space.



What is a 2-Prong Discriminant?

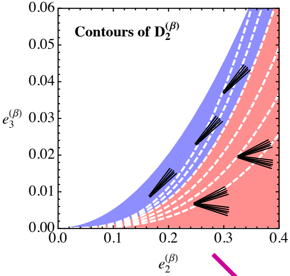
- Measure some set of observables on a jet, call them $e_2^{(\beta)}, e_3^{(\beta)}$.
- One and two prong jets live in distinct regions of the phase space.
- Define a family of contours, e.g. $e_3^{(\beta)} = D_2^{(\beta)} \left(e_2^{(\beta)}\right)^3$

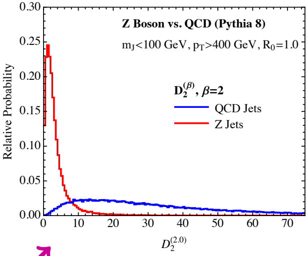


[See also *N*-subjettiness: Thaler, Van Tilburg]

What is a 2-Prong Discriminant?

- Form a real valued discriminant by marginalization:





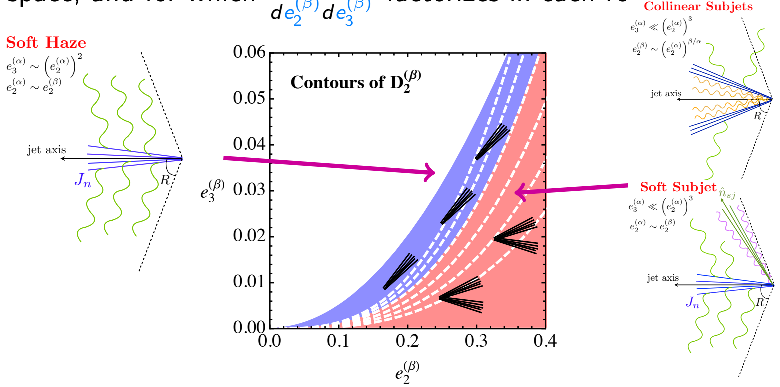
$$\frac{d\sigma}{dD_2^{(\beta)}} = \int de_2^{(\beta)} de_3^{(\beta)} \delta\left(D_2^{(\beta)} - \frac{e_3^{(\beta)}}{\left(e_2^{(\beta)}\right)^3}\right) \frac{d^2\sigma}{de_2^{(\beta)} de_3^{(\beta)}}$$

- D_2 is not IRC safe, but computable in resummed perturbation theory. [Larkoski, Thaler]

$$\frac{d\sigma}{dD_2^{(\beta)}} = \mathcal{O}\left(\sqrt{\alpha_s}\right) + \cdots$$

Factorizing a 2-Prong Discriminant

- Want to choose $e_2^{(\beta)}$, $e_3^{(\beta)}$ which parametrically separate phase space, and for which $\frac{d^2\sigma}{de_2^{(\beta)}de_3^{(\beta)}}$ factorizes in each region.



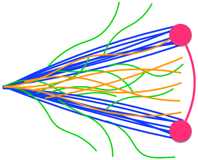
- Contours for a good discriminant must respect phase space separation

$\implies$ Marginalization can be performed in each EFT.

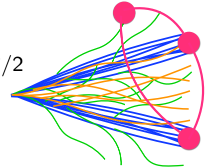
Choosing an Observable

- Use Energy Correlation Functions as basis of observables:
  [Larkoski, Salam, Thaler]

$$e_2^{(\beta)} = \frac{1}{E_J^2} \sum_{i<j\in J} E_i E_j \left(\frac{2p_i \cdot p_j}{E_i E_j} \right)^{\beta/2}$$



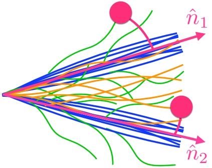
$$e_3^{(\beta)} = \frac{1}{E_J^3} \sum_{i<j<k\in J} E_i E_j E_k \left(\frac{2p_i \cdot p_j}{E_i E_j} \frac{2p_i \cdot p_k}{E_i E_k} \frac{2p_j \cdot p_k}{E_j E_k} \right)^{\beta/2}$$



- Defined without a fixed number of axes.

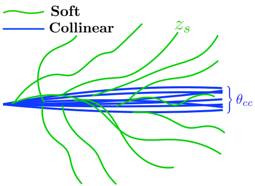
- Contrast with $\tau_N^{(\alpha)} = \frac{1}{E_J} \sum_{i\in J} E_i \min \{R_{i1}^\alpha, \dots, R_{iN}^\alpha\}$:

[Stewart, Tackmann, Waalewijn]
[Thaler, Van Tilburg]



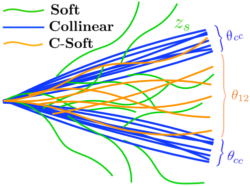
Power Counting: $e_2^{(\beta)}$, $e_3^{(\beta)}$ Phase Space

- Power counting determines structure of $e_2^{(\beta)}$, $e_3^{(\beta)}$ phase space:



$$e_2^{(\beta)} \sim \theta_{cc}^{\beta} + z_s\,,$$
$$e_3^{(\beta)} \sim \theta_{cc}^{3\beta} + z_s^2 + \theta_{cc}^{\beta} z_s$$
$$\Longrightarrow \text{ 1-prong jet: } (e_2^{(\beta)})^3 \lesssim e_3^{(\beta)} \lesssim (e_2^{(\beta)})^2$$

SCET$_+$:



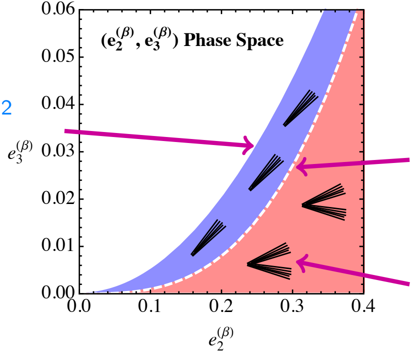
[Bauer, Tackmann, Walsh, Zuberi]

$$e_2^{(\beta)} \sim \theta_{12}^{\beta}\,,$$
$$e_3^{(\beta)} \sim \theta_{12}^{\beta} z_s + \theta_{12}^{2\beta}\theta_{cc}^{\beta} + \theta_{12}^{3\beta} z_{cs}$$
$$\Longrightarrow \text{ 2-prong jet: } 0 < e_3^{(\beta)} \ll (e_2^{(\beta)})^3$$

Power Counting: $e_2^{(\beta)}$, $e_3^{(\beta)}$ Phase Space

- Energy correlation functions parametrically separate phase space.
- Will show how to factorize in each region.

$$e_3^{(\beta)} = \frac{1}{2}\left(e_2^{(\beta)}\right)^2$$



$$e_3^{(\beta)} \sim \left(e_2^{(\beta)}\right)^3$$

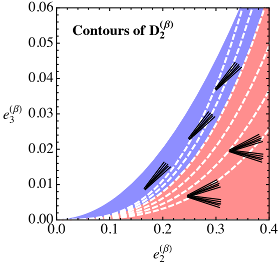
$$e_3^{(\beta)} \ll \left(e_2^{(\beta)}\right)^3$$

- Scaling of contours separating one from two prong jets is determined by power counting.

Choosing a Discriminating Observable: D_2

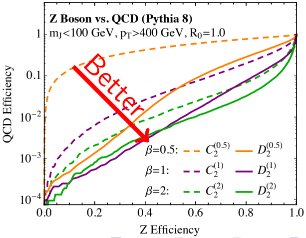
- Need to choose form of contours for optimal discrimination: power counting makes this trivial! No guess work.

$$D_2^{(\beta)} = \frac{e_3^{(\beta)}}{\left(e_2^{(\beta)}\right)^3} \qquad \left(\text{or } D_2^{(\alpha,\beta)} = \frac{e_3^{(\alpha)}}{(e_2^{(\beta)})^{3\alpha/\beta}}\right)$$



- Behaviour determined by parametrics $\implies$ robust to MC tuning.

- Improved performance in MC compared to other variables.



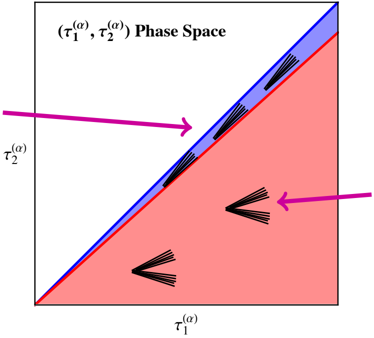
Aside: N-subjettiness

[Stewart, Tackmann, Waalewijn]
[Thaler, Van Tilburg]

- $\tau_1^{(\alpha)}, \tau_2^{(\alpha)}$ (N-subjettinesses) do not parametrically separate phase space.

$\tau_2^{(\alpha)} \sim \tau_1^{(\alpha)}$

Unknown Description



$$\tau_{2,1}^{(\alpha)} = \frac{\tau_2^{(\alpha)}}{\tau_1^{(\alpha)}}$$

$\tau_2^{(\alpha)} \ll \tau_1^{(\alpha)}$ ✓

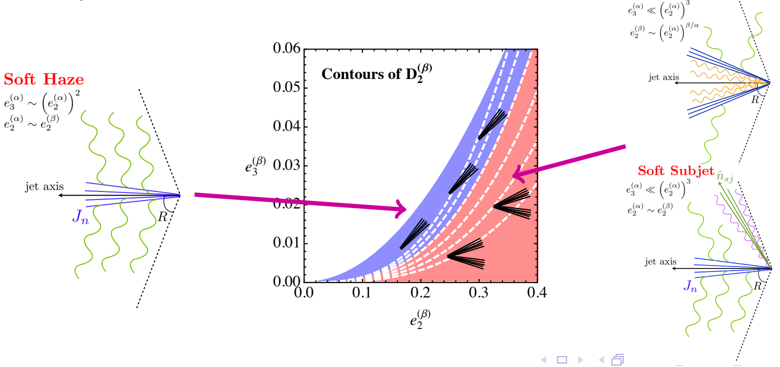
Studied in:

[Feige, Schwartz, Stewart, Thaler]

- One-prong region not straightforward. Non-singular matrix elements (boundary) important.

Factorization of D_2

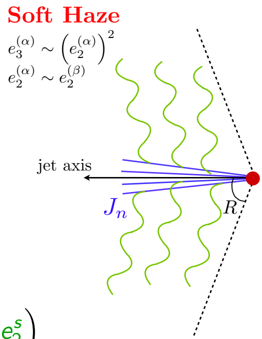
- Prove factorization theorem in each region of phase space: soft haze, collinear subjets, soft subjet.
    - Additional measurement required in two-prong region to distinguish collinear and soft subjets. Can be marginalized over.
- Factorized description of phase space regions can be combined to make prediction for D_2.



Factorization: Soft Haze

[Larkoski, IM, Neill 1504.soon]

- Only a single jet resolved by measurement of D_2.



- Factorization involving a single measured collinear sector:

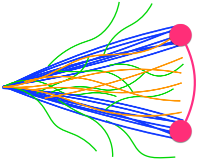
$$\frac{d\sigma}{de_2^{(\alpha)}de_3^{(\alpha)}} = H_{n\bar{n}}(Q^2)J_{\bar{n}} \int de_2^c de_2^s de_3^s \, \delta\left(e_2^{(\alpha)} - e_2^c - e_2^s\right) \delta\left(e_3^{(\alpha)} - e_2^c e_2^s - e_3^s\right) J_n\left(e_2^c\right) S_{n\bar{n}}\left(e_2^s, e_3^s\right)$$

- Non-trivial soft-collinear factorization of $e_3^{(\alpha)}$ observable.

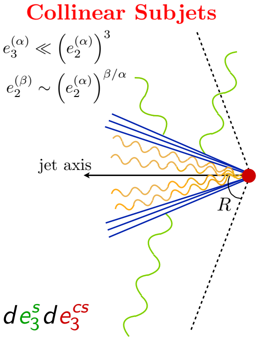
Factorization: Collinear Subjets (NINJA)

[Bauer, Tackmann, Walsh, Zuberi]

- Application of the SCET_+ EFT.
- $e_2^{(\beta)}$ set by hard splitting.



- Contributions to $e_3^{(\beta)}$ factorize: collinears, softs and collinear-softs.

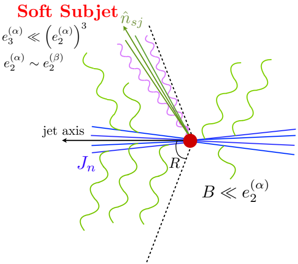


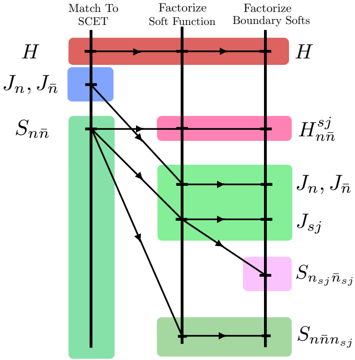
$$\frac{d\sigma}{dZ\, de_2^{(\alpha)}\, de_3^{(\alpha)}} = \sum_{f,f_a,f_b} H^f_{n_t\bar{n}_t} J_{\bar{n}_t} P^{f\to f_a f_b}_{n_t\to n_a,n_b}\left(Z; e_2^{(\alpha)}\right) \int de_3^c\, de_3^{\bar{c}}\, de_3^s\, de_3^{cs}$$

$$\delta\left(e_3^{(\alpha)} - e_3^c - e_3^{\bar{c}} - e_3^s - e_3^{cs}\right) J^{f_a}_{n_a}\left(Z; e_3^c\right) J^{f_b}_{n_b}\left(1-Z; e_3^{\bar{c}}\right) S_{n_t\bar{n}_t}(e_3^s) S^+_{n_a n_b \bar{n}_t}\left(e_3^{cs}\right)$$

[Larkoski, IM, Neill 1501.04596]

- Novel factorization theorem.
- "Boundary soft" modes resum NGLs associated with jet boundary.





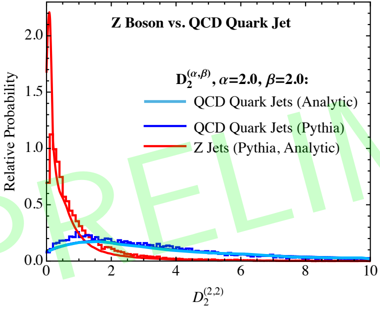
$$\frac{d\sigma(B;R)}{dz_{sj}\, de_2^{(\alpha)}\, de_3^{(\alpha)}} = H(Q^2) H_{n\bar{n}}^{sj}\left(z_{sj}, e_2^{(\alpha)}\right) J_n\left(e_3^{(\alpha)}\right) \otimes J_{\bar{n}}(B)$$

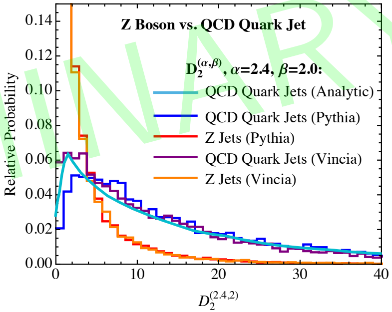
$$\otimes S_{n\bar{n}n_{sj}}\left(e_3^{(\alpha)}; B; R\right) \otimes J_{n_{sj}}\left(e_3^{(\alpha)}\right) \otimes S_{n_{sj}\bar{n}_{sj}}\left(e_3^{(\alpha)}; R\right)$$

(Preliminary) Comparison with Monte Carlo

[Larkoski, IM, Neill 1501.soon]

- Compare analytic D_2 prediction (NLL in each factorization theorem) with Parton level Monte Carlo.
- Use discrimination of $e^+e^- \to q\bar{q}$ and $e^+e^- \to ZZ \to q\bar{q}q\bar{q}$ at 1TeV as an example.

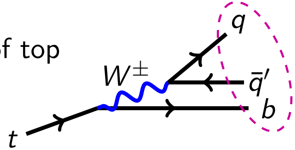




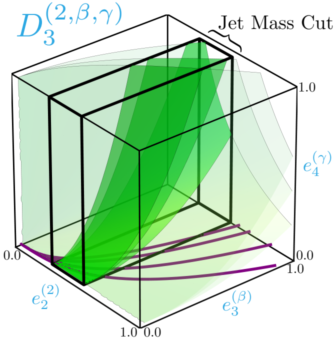
- Analytic boosted boson discrimination with a jet shape!
- Detailed probe of MC shower: Vincia(antenna) vs. Pythia(p_T ordered)

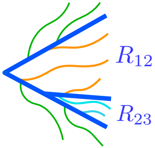
Boosted Top Taggers

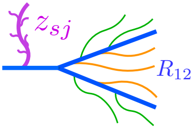
[Larkoski, IM, Neill 1411.0665]



- Approach can be extended to factorization of top taggers.
- Requires studying $e_2^{(\alpha)}$, $e_3^{(\beta)}$, $e_4^{(\gamma)}$ phase space.
- Power counting used to identify parametric scaling of surfaces separating 1,2, and 3 prong regions.
- Additional hierarchies present in description of 3 prong region.







Boosted Top Taggers

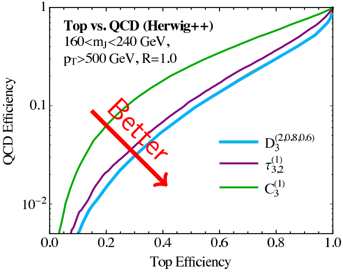
[Larkoski, IM, Neill 1411.0665]

- Detailed power counting analysis of $e_2^{(\alpha)}$, $e_3^{(\beta)}$, $e_4^{(\gamma)}$ phase space leads to:

$$D_3^{(\alpha,\beta,\gamma)} = \frac{e_4^{(\gamma)}\left(e_2^{(\alpha)}\right)^{\frac{3\gamma}{\alpha}}}{\left(e_3^{(\beta)}\right)^{\frac{3\gamma}{\beta}}} + x\frac{e_4^{(\gamma)}\left(e_2^{(\alpha)}\right)^{\frac{2\gamma}{\beta}-1}}{\left(e_3^{(\beta)}\right)^{\frac{2\gamma}{\beta}}} + y\frac{e_4^{(\gamma)}\left(e_2^{(\alpha)}\right)^{\frac{2\beta}{\alpha}-\frac{\gamma}{\alpha}}}{\left(e_3^{(\beta)}\right)^{2}}$$

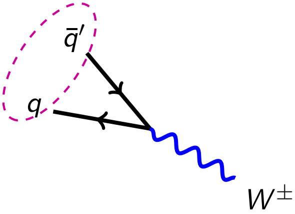
- Excellent performance in Monte Carlo.

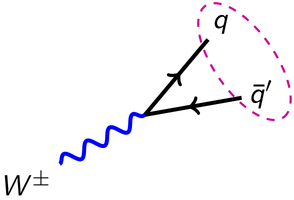
- Complete analytic calculation would provide considerable insight into top tagging observables, and ability of MC to describe them.



Conclusions

- Precision calculations of substructure observables important for LHC.
- Power counting is a powerful tool for understanding and designing substructure observables e.g. D_2, D_3
- Multi-differential factorization can bridge the gap between precision QCD calculations and substructure observables of phenomenological interest.





Thanks!