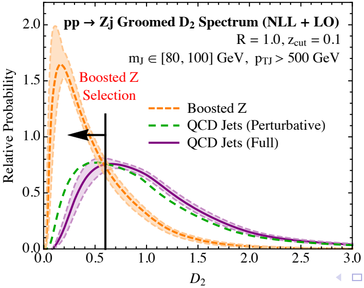
Analytic Boosted Boson Discrimination at the LHC

Ian Moult

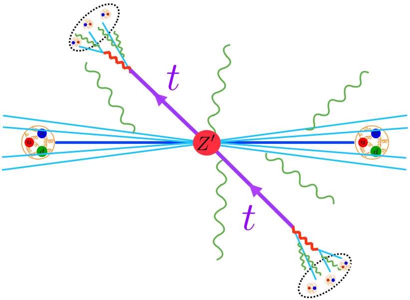
Berkeley



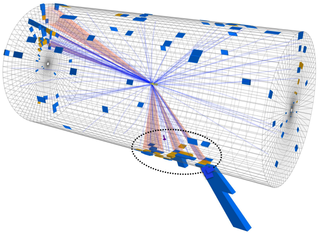
Jets at the LHC

- Internal structure of jets resolved due to excellent detector resolution.
- Electroweak scale objects, $W/Z/H$ or t can have sufficiently high p_T to appear inside a jet.

Boosted Tops



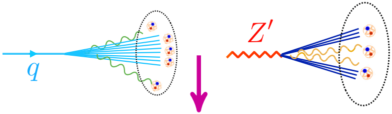
Event Display



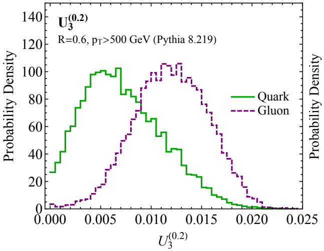
- Revolutionizes the types of questions we can/must ask about jets:
  $\Longrightarrow$ jets have substructure!

Jet Substructure

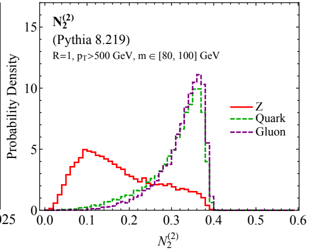
- Jets now act as proxies for almost all SM particles!
- Jet substructure: measure properties (charge, energy, etc) of radiation in a jet to extract information about its origin.



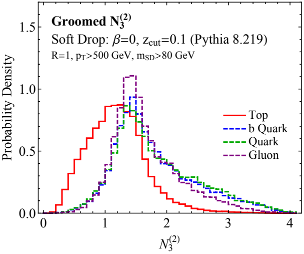
Quark vs. Gluon



$W/Z/H$ Tagging

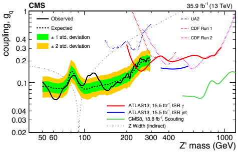


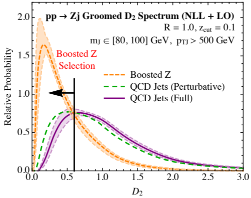
Top Tagging

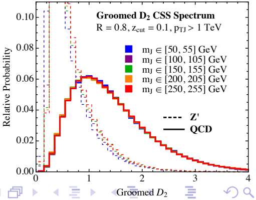


Outline

- Modern Observables and Applications
- Analytic Calculation of Groomed D_2
- Convolved Substructure

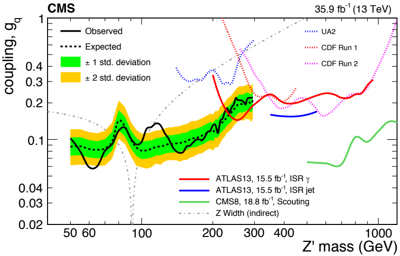






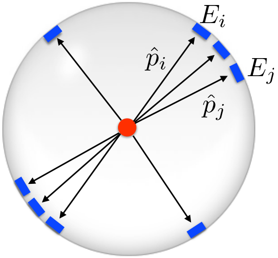
[Larkoski, IM, Neill]
[IM, Necib, Thaler]

Modern Observables and Applications



Substructure Observables

- A huge number of observables have been proposed over the years.
- Would like a common framework for jet substructure observables:



$$F_N(P) = \sum E_{i_1} \cdots E_{i_N} f_N(\hat{p}_{i_1}, \cdots, \hat{p}_{i_N})$$

- Linear in the energies by Infrared and Collinear (IRC) safety.
- f_N is symmetric, and $f_N \to 0$ if $\hat{p}_i || \hat{p}_j$

- Known that from this one can reconstruct any IRC safe observable.
- Is this useful for jet substructure?

$\Longrightarrow$ Need to choose a basis.

Generalized Energy Correlation Functions

$$R_{ij} = \sqrt{\Delta y_{ij}^2 + \Delta\phi_{ij}^2}$$

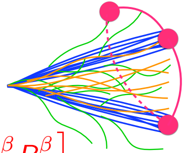
General Energy Correlation Functions

$$\boxed{{}_i e_j^{(\beta)} = \frac{1}{p_{TJ}^j} \sum_{1\le n_1<\cdots<n_j\le n} p_{Tn_1} p_{Tn_2} \dots p_{Tn_j} \min\left(\prod_{s,t}^{i} R_{st}^{\beta}\right)}$$

[IM, Necib, Thaler]

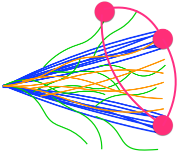
- Example: Three different ways to probe three particle correlations.

$$_1 e_3^{(\beta)} = \frac{1}{p_{TJ}^3} \sum_{1\le i<j<k\le n_J} p_{Ti} p_{Tj} p_{Tk} \min\left[R_{ij}^{\beta}, R_{ik}^{\beta}, R_{jk}^{\beta}\right],$$



$$_2 e_3^{(\beta)} = \frac{1}{p_{TJ}^3} \sum_{1\le i<j<k\le n_J} p_{Ti} p_{Tj} p_{Tk} \min\left[R_{ij}^{\beta} R_{ik}^{\beta}, R_{ij}^{\beta} R_{jk}^{\beta}, R_{ik}^{\beta} R_{jk}^{\beta}\right],$$

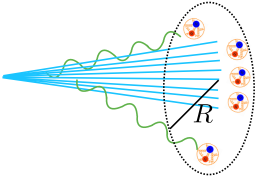
$$_3 e_3^{(\beta)} = \frac{1}{p_{TJ}^3} \sum_{1\le i<j<k\le n_J} p_{Ti} p_{Tj} p_{Tk} R_{ij}^{\beta} R_{ik}^{\beta} R_{jk}^{\beta} = e_3^{(\beta)}$$

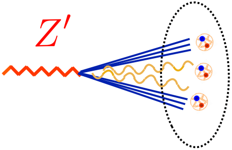


- Flexible basis for substructure observables.

Power Counting Observables

- How can we combine these observables to identify features of a jet?
- Example: Boosted $W/Z/H$ discrimination:

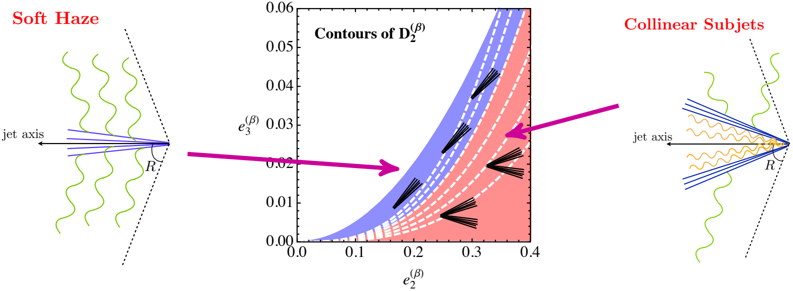




1. Write down Effective Field Theory (EFT) description of each configuration.
2. Identify region of validity of EFTs in terms of observables ¡$e_j^{(\beta)}$.
3. Shared boundaries define how to separate.

Power Counting Observables: D_2

- Consider using $e_2^{(\beta)}$, $e_3^{(\beta)}$.
- Phase space separated by contours of $D_2^{(\beta)}$: $e_3^{(\beta)} = D_2^{(\beta)} \left(e_2^{(\beta)} \right)^3$



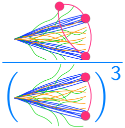
$\left(e_2^{(\beta)} \right)^3 < e_3^{(\beta)} < \left(e_2^{(\beta)} \right)^2$

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

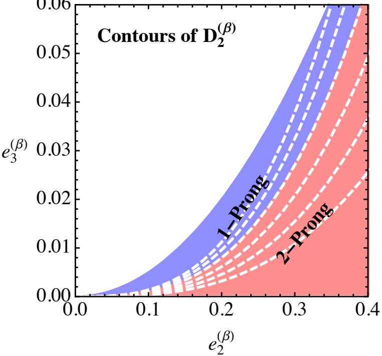
Phase Space

- Define the discriminant

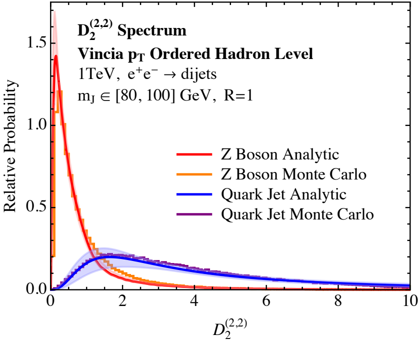
$$D_2^{(\beta)} = \frac{e_3^{(\beta)}}{\left(e_2^{(\beta)}\right)^3} =$$



$\left(e_2^{(\beta)}, e_3^{(\beta)}\right)$ Phase Space



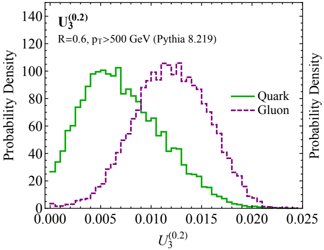
D_2 Spectrum



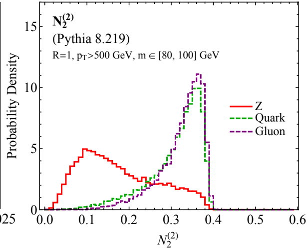
Summarizing the Observables

- D_2 part of a larger family of observables that can be designed for a variety of purposes.

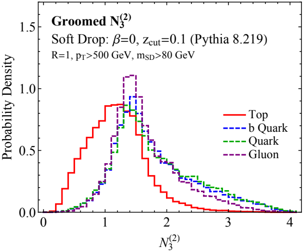
Quark vs. Gluon



$W/Z/H$ Tagging



Top Tagging

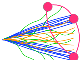


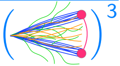
- Focus on $W/Z/H$:

Searches at ATLAS

[Larkoski, IM, Neill]

$$D_2^{(\beta)} = \frac{e_3^{(\beta)}}{\left(e_2^{(\beta)}\right)^3} =$$

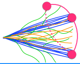


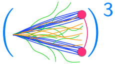


Searches at CMS

[IM, Necib, Thaler]

$$N_2^{(\beta)} = \frac{{}_2e_3^{(\beta)}}{(e_2^{(\beta)})^2}$$



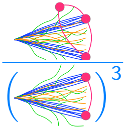


The Shape of Jets at the LHC: D_2

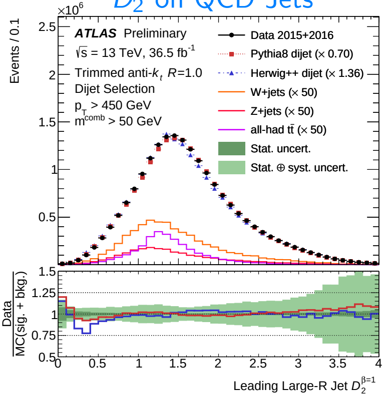
[Larkoski, IM, Neill]

- D_2 is default tagger for ATLAS.

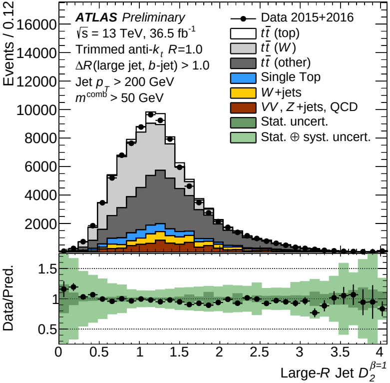
$$D_2^{(\beta)} = \frac{e_3^{(\beta)}}{\left(e_2^{(\beta)}\right)^3} =$$



D_2 on QCD Jets



D_2 on W Jets

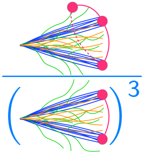


The Shape of Jets at the LHC: N_2

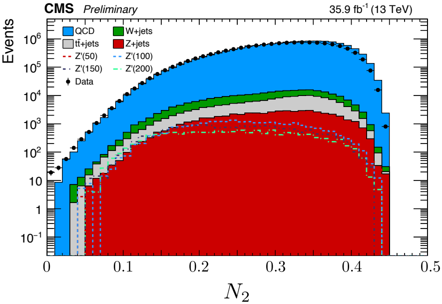
[IM, Necib, Thaler]

- N_2 used by CMS.

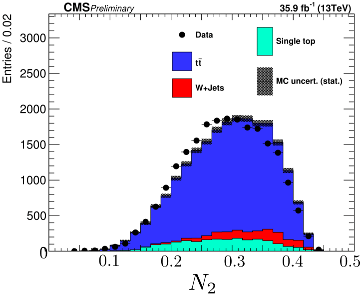
$$N_2^{(\beta)} = \frac{2\, e_3^{(\beta)}}{(e_2^{(\beta)})^2}$$



N_2 on QCD Jets

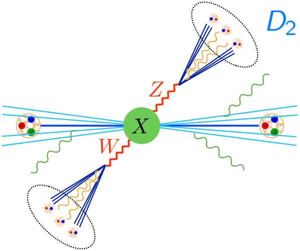


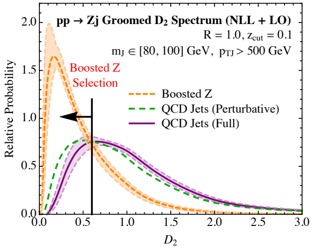
N_2 on W Jets



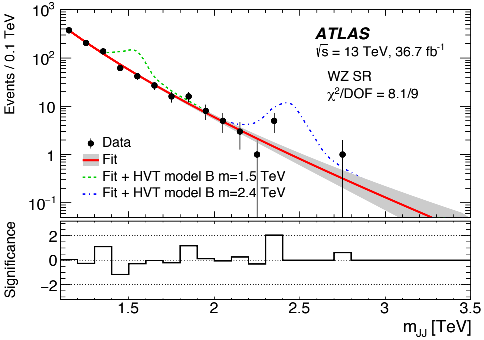
Canonical Applications: High Mass Resonances

- Canonical substructure application: Heavy resonance search
- Substructure only used as a tagger: “yes/no”





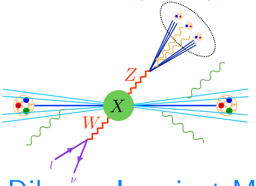
WZ Diboson Invariant Mass

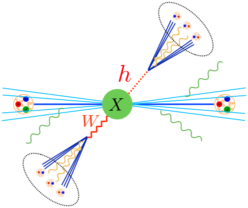


Using [Larkoski, IM, Neill]

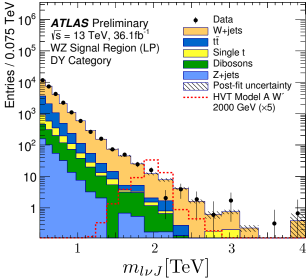
Canonical Applications: High Mass Resonances

- Particular interest in possibility of high mass resonances decaying to dibosons: WW, WZ, HZ, $\gamma\gamma$, γZ, $\cdots$

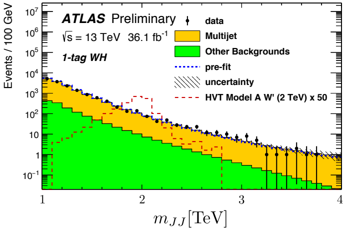




WZ Diboson Invariant Mass



WH Diboson Invariant Mass

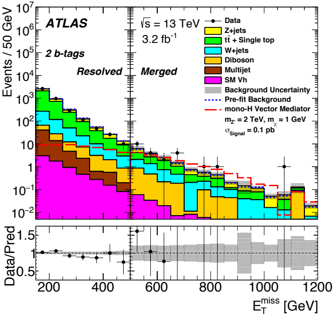


Using [Larkoski, IM, Neill]

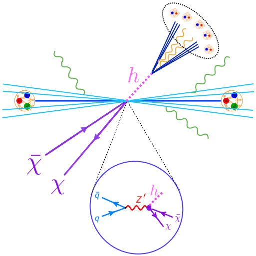
Canonical Applications: Dark Matter Searches

- Dark Matter Searches:

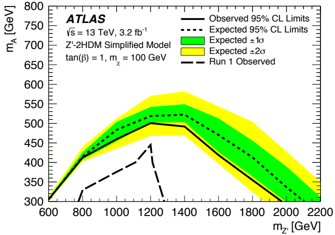
E_T^{miss}



Substructure



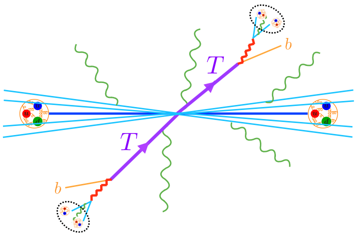
Mediator Limit



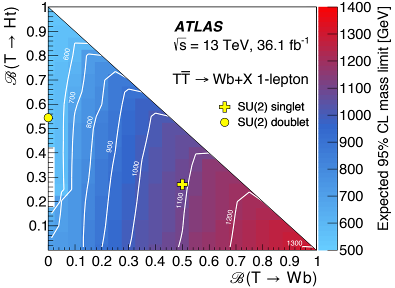
Using [Larkoski, IM, Neill]

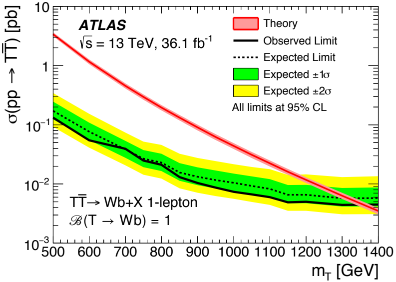
Applications: Vector Like Quark Searches

- Searches for vector like quarks:
  $pp \to TT \to WbWb$



Limits on VLQ Mass

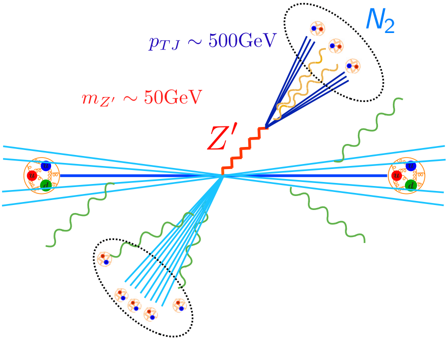




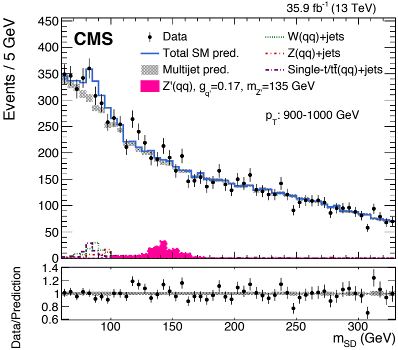
Using [Larkoski, IM, Neill]

Next Generation of Substructure Searches

- Beyond tagging: Look at distributions measured on a tagged jet
- Rely on detailed behavior of the QCD aspects of the substructure observables!



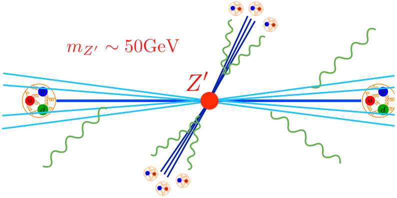
Tagged Jet Mass



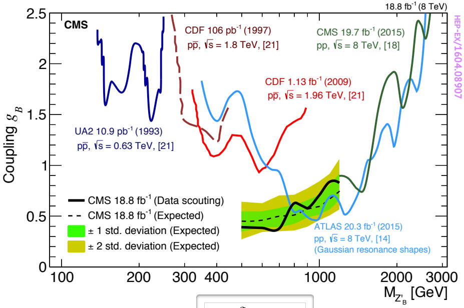
Using [IM, Necib, Thaler]

Low Mass Searches

- Why is this useful? Low masses (scales) hard to probe at LHC due to triggers, backgrounds, etc.



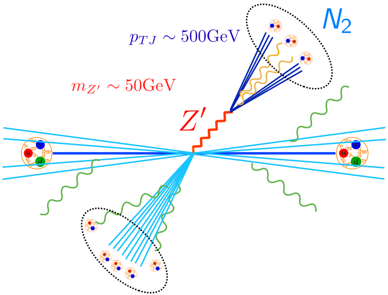
Limits on Light Z'



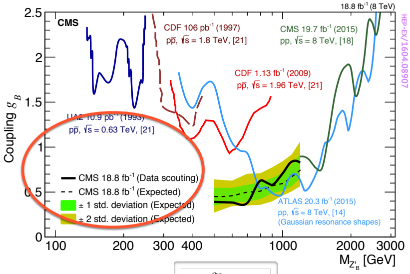
$$\mathcal{L} \sim \frac{g_B}{6} Z'_{B,\mu} \bar{q} \gamma^\mu q$$

Low Mass Searches

- Jet substructure can probe low scales within a high energy jet.
- Interesting region of parameter space currently unconstrained.

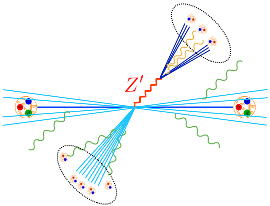


Prior Limits



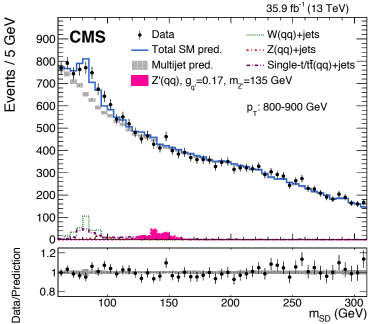
$$\mathcal{L} \sim \frac{g_B}{6} Z'_{B,\mu} \bar{q} \gamma^\mu q$$

Low Mass Searches

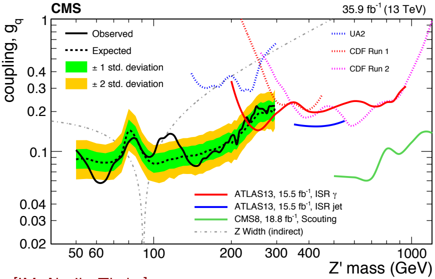


- Low mass Z' search in dijet channel.

Jet Mass Distribution



Limits from N_2



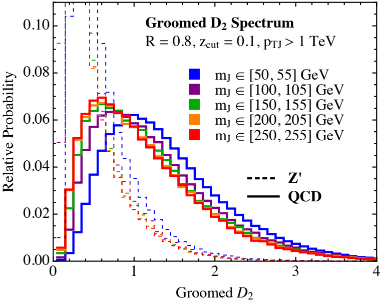
Using [IM, Necib, Thaler]

- Find $Z' = Z/W$!
- Probe new low mass, low x-sec region of parameter space!

Modern Substructure Analyses

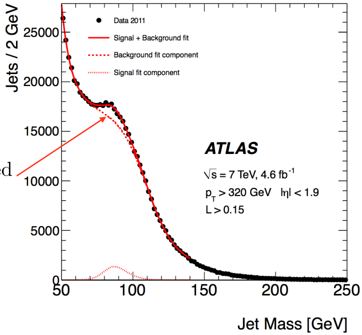
- Highly non-trivial from QCD perspective!
- Requires improved understanding of substructure observables.

D_2 Mass Dependence

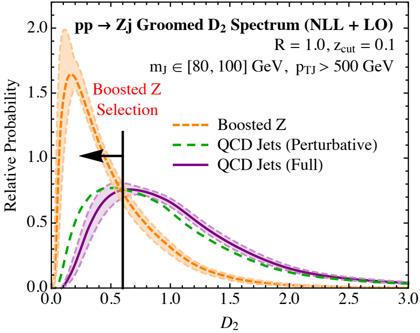


Mass Spectrum After Cut

Background Shaped by Cuts!



Analytic Calculation of Groomed D_2

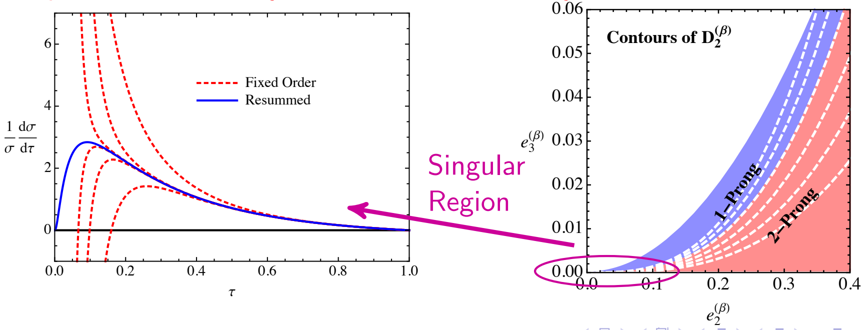


D_2

- Cross section for D_2 computed by marginalization:

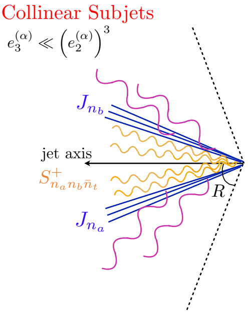
$$\frac{d\sigma}{dD_2} = \int de_2^{(\alpha)} de_3^{(\alpha)}\, \delta\left(D_2 - \frac{e_3^{(\alpha)}}{(e_2^{(\alpha)})^3}\right) \frac{d\sigma}{de_2^{(\alpha)} de_3^{(\alpha)}}$$

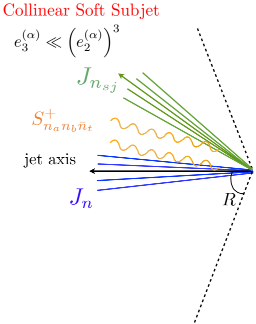
- For each value of D_2 contour of integration passes through singular region of phase space $\Longrightarrow$ not computable in fixed order perturbation theory, resummation necessary!



EFTs for Jet Substructure

- Can understand all orders structure using Effective Field Theories.
- Dynamics of subjets iteratively integrated out and replaced by sources (Wilson lines)

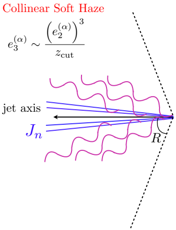


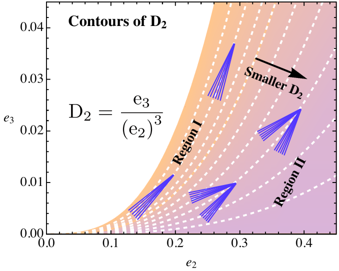


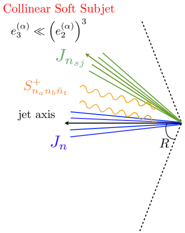
- Powerful approach to describe complicated situations.
- Extensions of SCET involving additional modes/ hierarchies.

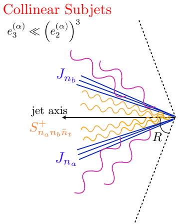
EFTs for Jet Substructure

- Tile the multi-differential phase space with EFTs





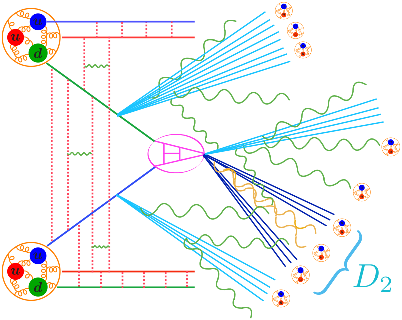




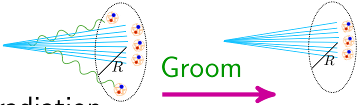
- At NLO involves both:
  - Single emission off two resolved sources.
  - Three unresolved particles.

Analytic Boosted Boson Discrimination at the LHC

- Difficulties in substructure calculations for *pp*:



- Global color correlations
- Hadronization corrections
- Pile-Up
- Underlying event



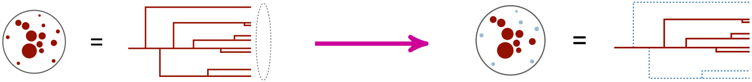
- All complications associated with soft radiation.
- Groomers remove soft radiation

  $\Longrightarrow$ Makes calculations simpler and more universal.

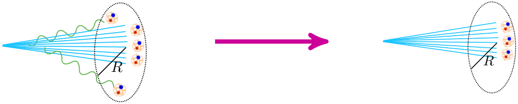
Soft Drop Grooming

[Larkoski, Marzani, Soyez, Thaler]

- Experimentally, a groomer is used to remove soft contamination.
- Soft Drop: Recurse through a Cambridge-Aachen clustering tree and remove particles that fail the condition: $\dfrac{\min[p_{Ti}, p_{Tj}]}{p_{Ti} + p_{Tj}} > z_{\mathrm{cut}}$



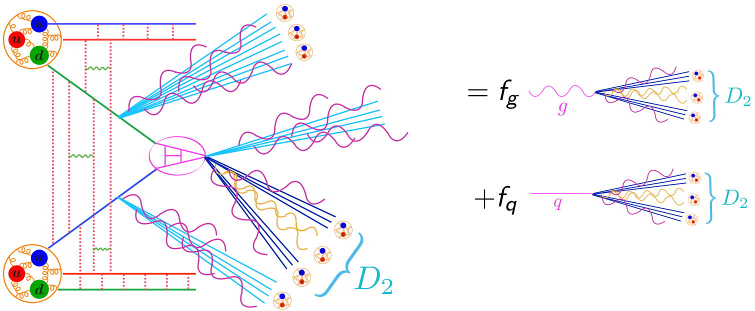
- Loosely speaking, reduces a jet to its collinear core.



- Any IRC safe observable measured on a groomed jet is IRC safe.

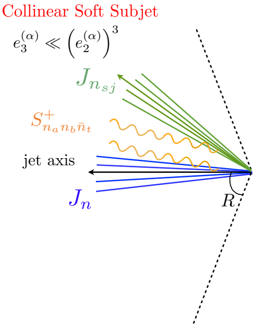
Analytic Boosted Boson Discrimination at the LHC

- (It can be shown that) Grooming removes all color correlations.



- Jet can be considered in isolation!
- Enables calculations in complicated LHC environment.

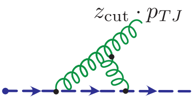
Perturbative Behavior

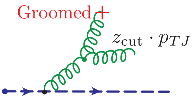


- Perturbative behavior is complicated.
- Matching coefficient to two-Wilson line operators exhibits interesting behavior.

$$\frac{d^2\sigma_k^{\mathrm{IIb}}}{dz\, de_2\, de_3} = H^s(z, e_2, z_{\mathrm{cut}})\, C_s(e_3)\, J_{sc}(e_3)\, J(e_3)$$

- "Groomed Soft Current":



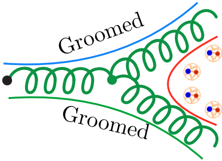


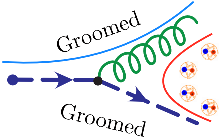
$$\gamma_{H^s} = -\left(2C_F + C_A\right)\Gamma_{\mathrm{cusp}}[\alpha_s]\log\frac{4\mu^2}{ze_2 Q^2} - \frac{\alpha_s}{2\pi}\beta_0 - \frac{\alpha_s C_F}{\pi}\log\frac{z^2}{z_{\mathrm{cut}}^2} + \frac{\alpha_s}{\pi}\left(C_F - C_A\right)\frac{\mathrm{Cl}_2\left(\frac{\pi}{3}\right)}{\pi}$$

- Will appear generally for groomed multi-prong observables.

Non-Perturbative Behavior

- Jet observables also receive non-perturbative corrections.
  - Hadronization
  - Underlying event
- For groomed two-prong observables, these are a property of the jet itself, and arise only from hadronization!
- In pictures:



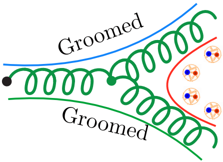


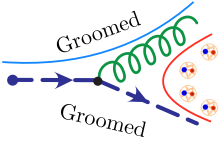
- As an operator statement:

$$C_{si}(e_3) = \mathrm{tr}\langle 0|\, T\{Y_i\}\delta(e_3 - \hat{\mathbf{E}}_3)\Theta_{\mathrm{SD}}\, \bar{T}\{Y_i\}|0\rangle$$

Non-Perturbative Behavior

- Non-perturbative effects exhibit a number of remarkable features
  - Negligible contribution from MPI/Underlying Event
  - Non-perturbative power corrections are suppressed by the jet mass
  - Independent of quark or gluon nature of jet.



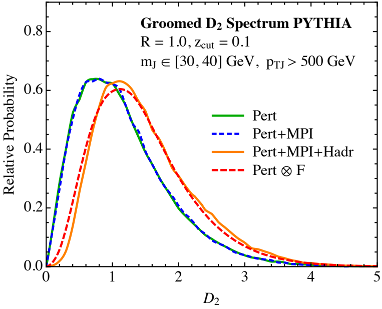


- Hadronization corrections for all jet masses described by a single non-perturbative parameter that can be extracted from e^+e^-!

Non-Perturbative Behavior

- Contribution from MPI/Underlying Event completely negligible.
- Non-perturbative corrections are from hadronization within the jet.

Groomed D2



- Non-perturbative corrections depends only on quark/gluon fraction, jet mass, and p_T.

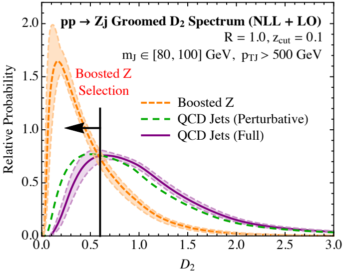
Non-Perturbative Behavior

- Hadronization corrections incorporated via a shape function $F(\epsilon)$.

$$\frac{d\sigma_{\text{NP}}}{dD_2} = \int_0^\infty d\epsilon\, F(\epsilon) \frac{d\sigma}{dD_2}\left(D_2 - \frac{\epsilon}{m_J z_{\text{cut}}^{3/2}}\right)$$

- Can be expanded in moments for $\Lambda_{\text{QCD}} \ll m_J D_2$
- Hadronization shifts the first moment of the D_2 distribution by

NP Shift



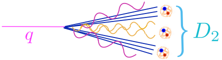
$$\Delta_D^{\text{NP}} = \frac{\Omega_D}{m_J z_{\text{cut}}^{3/2}}$$

- Here $\Omega_D \sim \Lambda_{\text{QCD}}$ is a universal parameter that can be extracted from e^+e^-.

Non-Perturbative Behavior

- Non-Perturbative correction controlled by perturbative jet mass

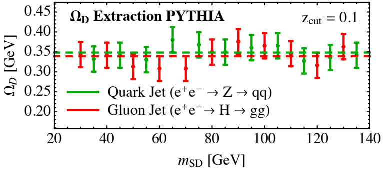
  $\Longrightarrow$ behaves like a (boosted) event shape with $Q = m_J$!



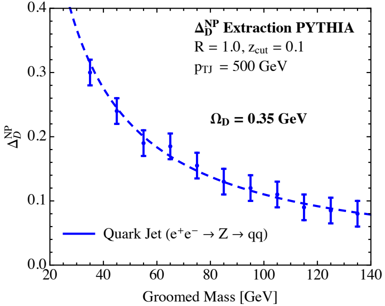
$$\Delta_D^{\mathrm{NP}} = \frac{\Omega_D}{m_J z_{\mathrm{cut}}^{3/2}}$$

$$\frac{d\sigma_{\mathrm{NP}}}{dD_2} = \int_0^\infty d\epsilon\, F(\epsilon) \frac{d\sigma}{dD_2}\left(D_2 - \frac{\epsilon}{m_J z_{\mathrm{cut}}^{3/2}}\right)$$

NP Parameter Extraction



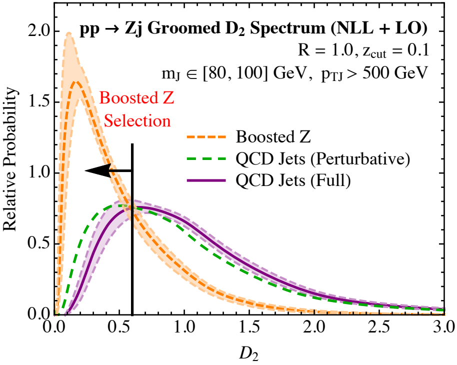
First Moment Shift



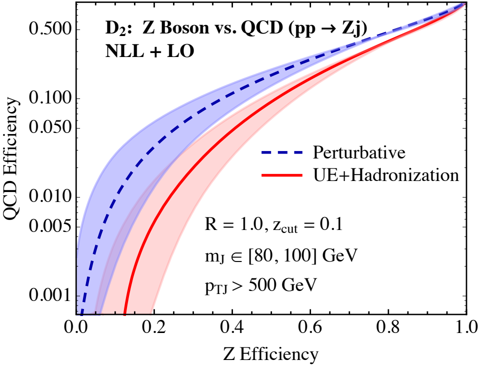
Analytic Boosted Boson Discrimination at the LHC

- Calculation of groomed D_2 at the LHC.

Groomed D_2 Distribution



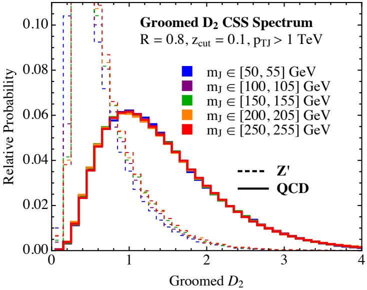
Discrimination Power



- Analytic understanding of modern jet substructure tools at LHC!

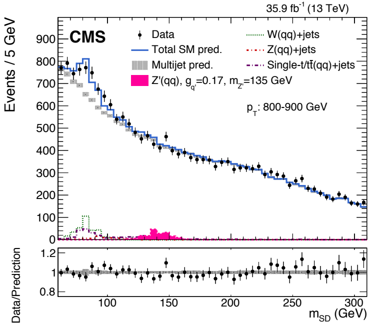
[IM, Nachmann, Neill]

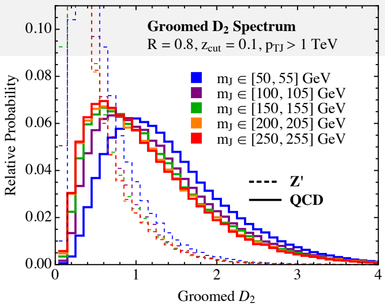
Convolved Substructure

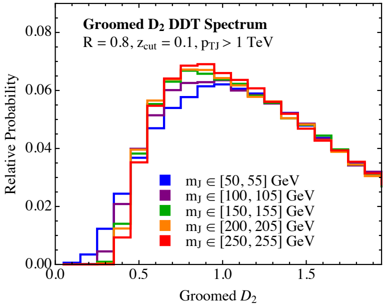


Low Mass Searches

- Stability improved using
  - Grooming
  - Decorrelation of first moment (DDT)
    [Dolen, Harris, Marzani, Rappoccio, Tran (CMS)]







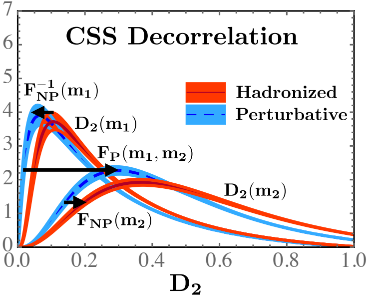
- Dependence on mass currently one of the primary difficulties.

Convolved Substructure

- Use derived factorization formula to decorrelate the D_2 distribution.
- Mass dependence of perturbative and non-perturbative components understood analytically.
- Use convolution with a "shape" function to take to a reference mass: Convolved Substructure (CSS)

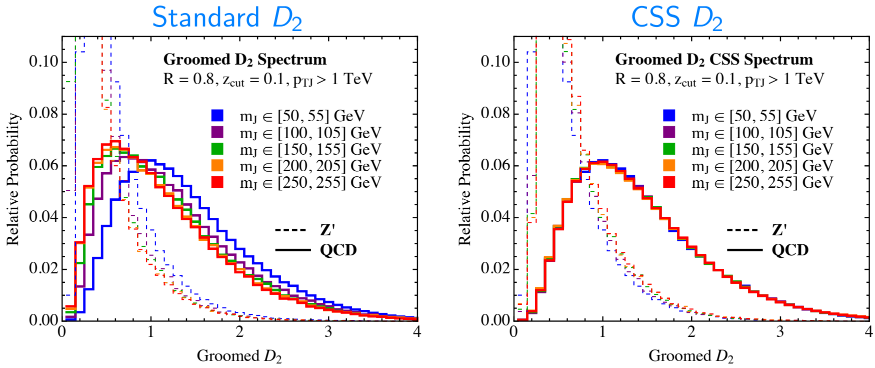
$$\frac{d\sigma^{\text{CSS}}}{dD_2} = \int_0^\infty d\epsilon\, F_{\text{CSS}}(\epsilon) \frac{d\sigma}{dD_2}(D_2 - \epsilon)$$

$$F_{\text{CSS}}(m_1, m_2) = F_{\text{NP}}^{-1}(m_1) \otimes F_{\text{P}}(m_1, m_2) \otimes F_{\text{NP}}(m_2)$$



Convolved Substructure: $Z' \to q\bar{q}$

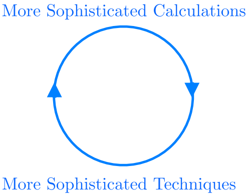
- Test CSS for light $Z' \to q\bar{q}$ search.



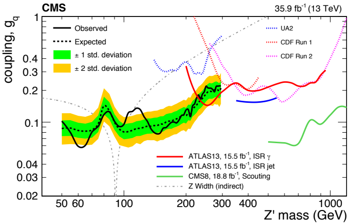
- Complete decorrelation of shape over wide range of masses.
- Currently being tested by ATLAS and CMS.
- Exploits detailed pert + non-pert behavior.

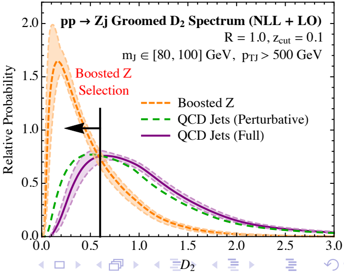
Conclusions

- Jet substructure provides novel ways to test the SM and to search for new physics at the LHC.



- Analytic calculations are a catalyst for further improvements in jet substructure.
- Substructure entering next level of sophistication.





Conclusions

- Shameless plug of our review on jet substructure:

Jet Substructure at the Large Hadron Collider:
A Review of Recent Advances in Theory and Machine Learning

Andrew J. Larkoski*
Physics Department, Reed College, Portland, OR 97202, USA

Ian Moult†
Berkeley Center for Theoretical Physics, University of California, Berkeley, CA 94720, USA and
Theoretical Physics Group, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA

Benjamin Nachman‡
Physics Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA
(Dated: September 15, 2017)

Jet substructure has emerged to play a central role at the Large Hadron Collider (LHC), where it has provided numerous innovative new ways to search for new physics and to probe the Standard Model in extreme regions of phase space. In this article we provide a comprehensive review of state of the art theoretical and machine learning developments in jet substructure. This article is meant both as a pedagogical introduction, covering the key physical principles underlying the calculation of jet substructure observables, the development of new observables, and cutting edge machine learning techniques for jet substructure, as well as a comprehensive reference for experts. We hope that it will prove a useful introduction to the exciting and rapidly developing field of jet substructure at the LHC.

Thanks!