Analytic Boosted Boson Discrimination

Ian Moult

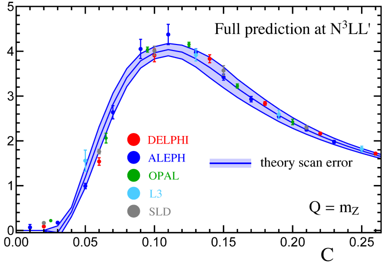
Massachusetts Institute of Technology

Based on 1507.03018 with Andrew Larkoski and Duff Neill

Motivation

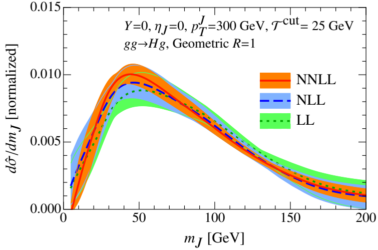
- Inclusive properties of events (thrust, C-parameter,...) and jets (mass, angularities,...) "well" understood:
  - All orders factorization theorems (systematically improveable) ✓
  - Treatment of non-perturbative physics (universality) ✓
  - Resummation to high orders (NNLL or NNNLL) ✓
- Allows precision comparison to data and Monte Carlo programs.

[Hoang, Kolodrubetz, Mateu, Stewart]



C-Parameter in e^+e^-

[Jouttenus, Stewart, Tackmann, Waalewijn]

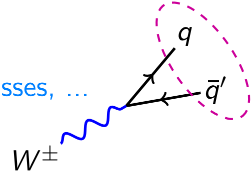


Jet Mass in $gg \to Hg$

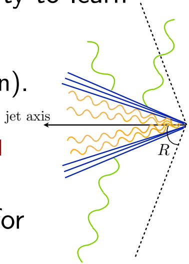
Motivation

- Main focus of last few years (and this conference) has been on jet substructure observables

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



- Pose new theoretical difficulties and present opportunity to learn about QCD
  - Additional hierarchical scales (complicate resummation).
  - Non-Global logarithms. See Andrew Larkoski's Talk
  - Sudakov safety of ratio observables. [Larkoski, Marzani, Thaler]

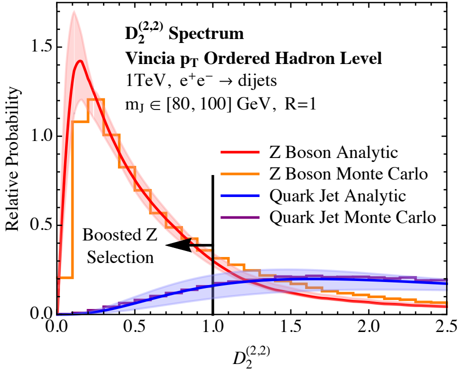


- Analytic calculations in jet substructure are essential for
  - Designing improved discrimination variables/ algorithms (mMDT, Soft Drop, ...)
  - Improving/understanding Monte Carlo description of the QCD shower

Goals

- Present an effective field theory framework for the calculation of 2-prong substructure observables. Relevant for W/Z/H tagging.
- Explicit calculation for a 2-prong discriminant D_2.

- All orders factorization theorems ✓
- Treatment of non-perturbative physics ✓
- Resummation to first non-trivial order (NLL) ✓



Approach to Calculation

Experiment

- Measure collection of IRC safe observables: $\tau_1^{(\beta)}, e_2^{(\beta)}, \cdots$
- Impose cuts on observables to classify different jet structures.
- Events in each classification separately treated.

Approach to Calculation

Experiment

- Measure collection of IRC safe observables: $\tau_1^{(\beta)}$, $e_2^{(\beta)}, \cdots$
- Impose cuts on observables to classify different jet structures.
- Events in each classification separately treated.

Calculation

- Calculate collection of IRC safe observables: $\tau_1^{(\beta)}$, $e_2^{(\beta)}, \cdots$
- Parametric relations between observables define classification.
- Effective field theory description of jets in each classification used for calculation.

Approach to Calculation

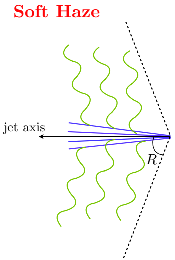
Experiment

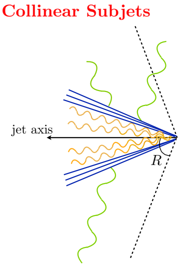
- Measure collection of IRC safe observables: $\tau_1^{(\beta)}, e_2^{(\beta)}, \cdots$
- Impose cuts on observables to classify different jet structures.
- Events in each classification separately treated.

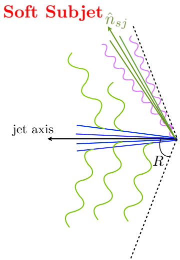
Calculation

- Calculate collection of IRC safe observables: $\tau_1^{(\beta)}, e_2^{(\beta)}, \cdots$
- Parametric relations between observables define classification.
- Effective field theory description of jets in each classification used for calculation.

EFTs for 2-prong Substructure:



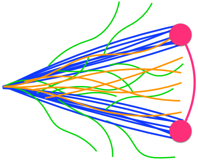




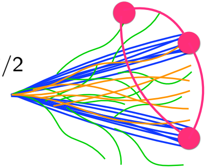
Choosing an Observable

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

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



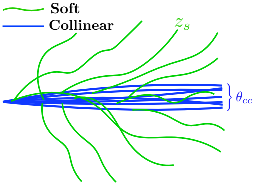
$$e_3^{(\beta)} = \frac{1}{E_J^3} \sum_{i<j<k\in J} E_i E_j E_k \left(\frac{2p_i \cdot p_j}{E_i E_j} \frac{2p_i \cdot p_k}{E_i E_k} \frac{2p_j \cdot p_k}{E_j E_k} \right)^{\beta/2}$$



- Sensitive to two-prong substructure.
- Simple structure. No minimization, no regions, behave as an event shape.

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

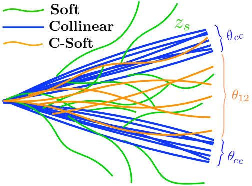
- Power counting determines structure of $e_2^{(\beta)}$, $e_3^{(\beta)}$ phase space:



$$e_2^{(\beta)} \sim \theta_{cc}^{\beta} + z_s\,,$$

$$e_3^{(\beta)} \sim \theta_{cc}^{3\beta} + z_s^2 + \theta_{cc}^{\beta} z_s$$

$$\Longrightarrow \text{ 1-prong jet: } (e_2^{(\beta)})^3 \lesssim e_3^{(\beta)} \lesssim (e_2^{(\beta)})^2$$



$$e_2^{(\beta)} \sim \theta_{12}^{\beta}\,,$$

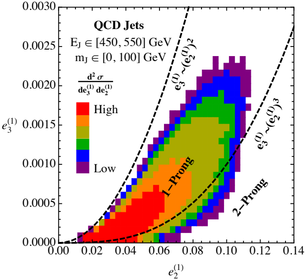
$$e_3^{(\beta)} \sim \theta_{12}^{\beta} z_s + \theta_{12}^{2\beta} \theta_{cc}^{\beta} + \theta_{12}^{3\beta} z_{cs}$$

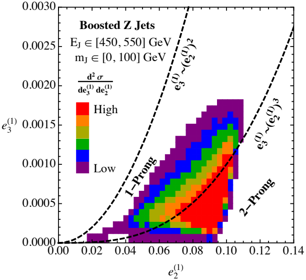
$$\Longrightarrow \text{ 2-prong jet: } 0 < e_3^{(\beta)} \ll (e_2^{(\beta)})^3$$

See Larkoski Boost 2014

Phase Space

- Parametric scaling dominates behavior of Monte Carlo

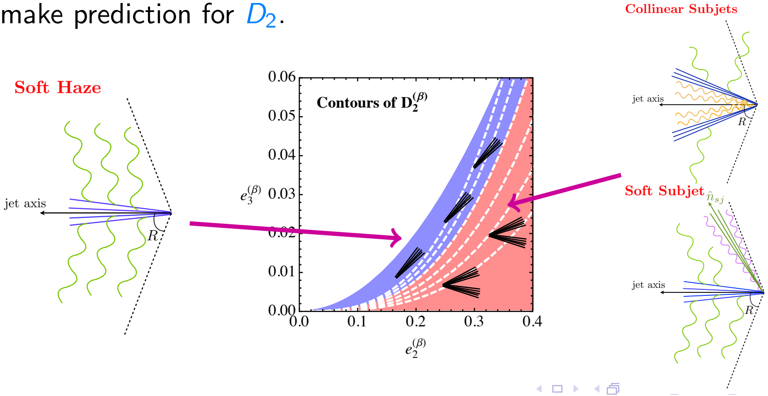




- $D_2^{(\beta)} = \frac{e_3^{(\beta)}}{\left(e_2^{(\beta)}\right)^3}$ $\left(\text{or } D_2^{(\alpha,\beta)} = \frac{e_3^{(\alpha)}}{(e_2^{(\beta)})^{3\alpha/\beta}}\right)$ provides a powerful variable for two-prong discrimination.
- Theoretically, D_2 allows for a classification of 1 vs 2 prong jets.

Factorization of D_2

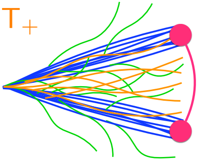
- Set up EFT description in each region of phase space: soft haze, collinear subjets, soft subjet.
  - Additional measurement required in two-prong region to distinguish collinear and soft subjets. Can be marginalized over.
- Factorized description of phase space regions can be combined to make prediction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



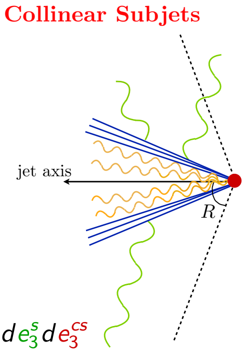
Factorization: Collinear Subjets

[Bauer, Tackmann, Walsh, Zuberi]

- Effective field theory description of radiation from collinear subjets: SCET$_+$
- $e_2^{(\beta)}$ set by hard splitting.



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



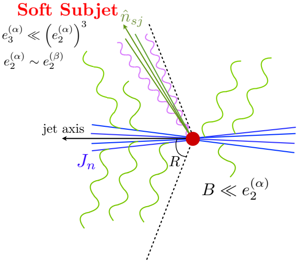
$$\frac{d\sigma}{dZ\, de_2^{(\alpha)}\, de_3^{(\alpha)}} = \sum_{f,f_a,f_b} H^f_{n_t \bar{n}_t} J_{\bar{n}_t} P^{f\to f_a f_b}_{n_t \to n_a, n_b}\left(Z; e_2^{(\alpha)}\right) \int de_3^c\, de_3^{\bar{c}}\, de_3^s\, de_3^{cs}$$

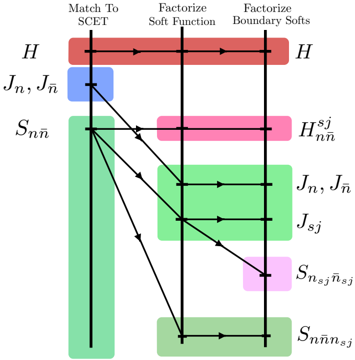
$$\delta\left(e_3^{(\alpha)} - e_3^c - e_3^{\bar{c}} - e_3^s - e_3^{cs}\right) J^{f_a}_{n_a}\left(Z; e_3^c\right) J^{f_b}_{n_b}\left(1 - Z; e_3^{\bar{c}}\right) S_{n_t \bar{n}_t}(e_3^s) S^+_{n_a n_b \bar{n}_t}\left(e_3^{cs}\right)$$

Factorization: Soft Subjet

[Larkoski, IM, Neill 1501.04596]

- Effective field theory description of radiation from a soft subjet:
- "Boundary soft" modes sensitive to jet boundary.





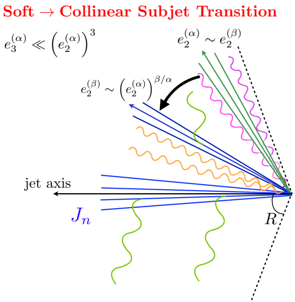
$$\frac{d\sigma(B;R)}{dz_{sj}\, de_2^{(\alpha)}\, de_3^{(\alpha)}} = H(Q^2)H_{n\bar{n}}^{sj}\left(z_{sj}, e_2^{(\alpha)}\right) J_n\left(e_3^{(\alpha)}\right) \otimes J_{\bar{n}}(B)$$

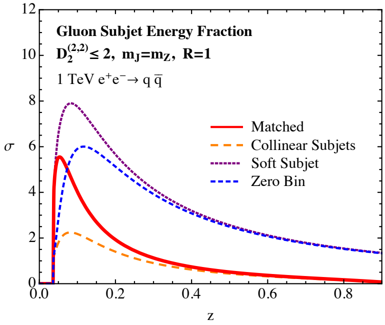
$$\otimes S_{n\bar{n}n_{sj}}\left(e_3^{(\alpha)}; B; R\right) \otimes J_{n_{sj}}\left(e_3^{(\alpha)}\right) \otimes S_{n_{sj}\bar{n}_{sj}}\left(e_3^{(\alpha)}; R\right)$$

Soft → Collinear Transition

- Soft subjet and collinear subjets can be merged to provide a complete description of the Soft → Collinear Subjet transition
- Zero bin procedure used to subtract overlap in description

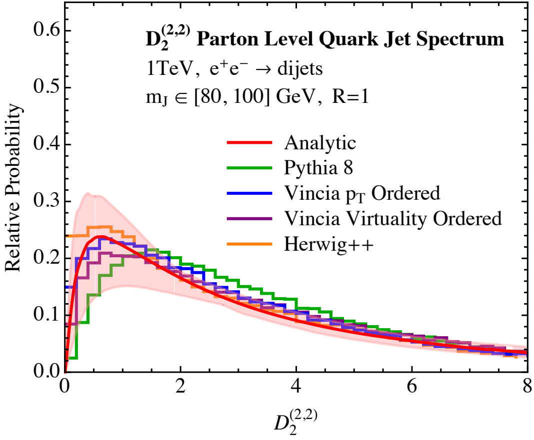
$$\sigma = (\sigma_{sj} - \sigma_{sj}|_{cs}) + \sigma_{cs}$$





Perturbative Comparisons

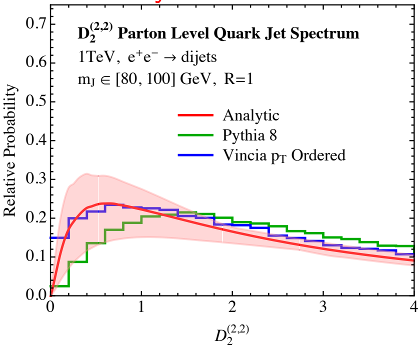
- D_2 is a sensitive probe of the perturbative shower: probes 3 particle correlations, and detailed structure of $1 \to 2$ splittings.



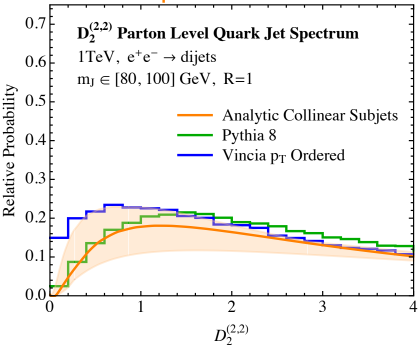
Perturbative Comparisons: Vincia vs. Pythia

- Understand impact of the soft subjet region of phase space:
  - Pythia: $1 \to 2$ collinear splittings (p_T ordering)
  - Vincia: $2 \to 3$ dipole antenna shower. Correct description of soft wide angle radiation (soft eikonal factor)

Analytic Calculation



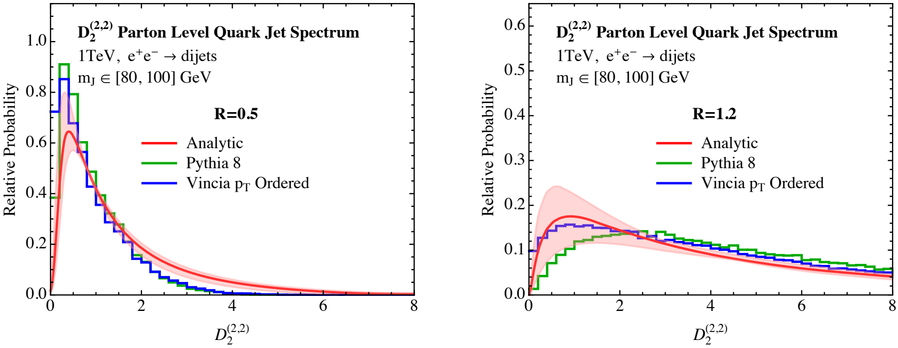
Incomplete Calculation



- Parton level behavior of Herwig (angular ordering) similar to Vincia.

Perturbative Comparisons: R dependence

- R dependence allows for further test of importance of soft subjet.

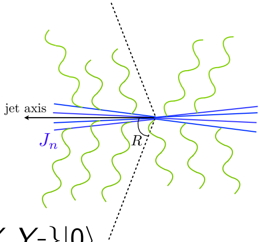


- Excellent agreement between analytic calculation and MC over a wide range of R values.

Non-Perturbative Effects

- Observables which are sensitive to soft radiation are sensitive to non-perturbative effects (hadronization).
- Implemented in MC with models. e.g. String or cluster fragmentation.
- All orders factorization theorem allows for a separation of perturbative and non-perturbative physics: soft function $S(\mathcal{M})$, describes (non-) perturbative soft radiation between jets

$$S(\mathcal{M}) = \int_0^\infty d\epsilon\, S_{\text{pert}}(\mathcal{M}-\epsilon)F(\epsilon)$$



- $F(\epsilon)$ is a model shape function.
- Field theoretic definition

$$S(\mathcal{M}) = \frac{1}{N_C}\mathrm{Tr}\langle 0|\,\bar{T}\{Y_{\bar{n}}^\dagger Y_n^\dagger\}\delta(\widehat{\mathcal{M}}-\mathcal{M})T\{Y_n Y_{\bar{n}}\}|0\rangle$$

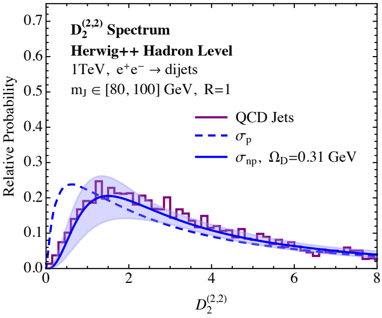
allows for relations between different observables. Maintain predictivity.

Non-Perturbative Effects for D_2

- Power counting a single non-perturbative emission from the subjets

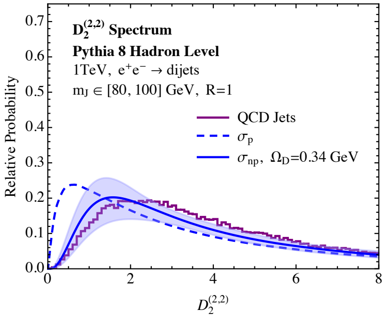
$$\frac{d\sigma_{\mathrm{np}}}{dD_2^{(\alpha,\beta)}} = \int_0^\infty d\epsilon\, F(\epsilon)\, \frac{d\sigma_{\mathrm{p}}\left(D_2^{(\alpha,\beta)} - \frac{\epsilon}{E_J}\frac{e_2^{(\alpha)}}{(e_2^{(\beta)})^{3\alpha/\beta}}\right)}{dD_2^{(\alpha,\beta)}}$$

- Consider $F(\epsilon) = \frac{4\epsilon}{\Omega_D{}^2} e^{-2\epsilon/\Omega_D}$
- Leading effect described by a single non-perturbative parameter $\Omega_D \sim \Lambda_{\mathrm{QCD}}$
- Ω_D fit to Monte Carlo:
    - Herwig (cluster): $\Omega_D = 0.31$ GeV
    - Vincia (string): $\Omega_D = 0.34$ GeV
    - Pythia (string): $\Omega_D = 0.47$ GeV



Hadron Level Comparisons

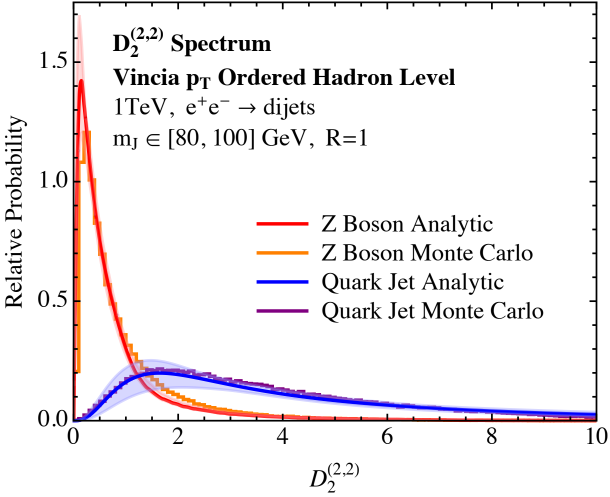
- Differences in parton level distributions decreased by hadronization, but not removed if a single $\Omega_D = 0.34$ GeV used for string model.



- Analytic perturbative and non-perturbative predictions allow to disentangle differences in perturbative and non-perturbative MC.
- Comparison with data essential!

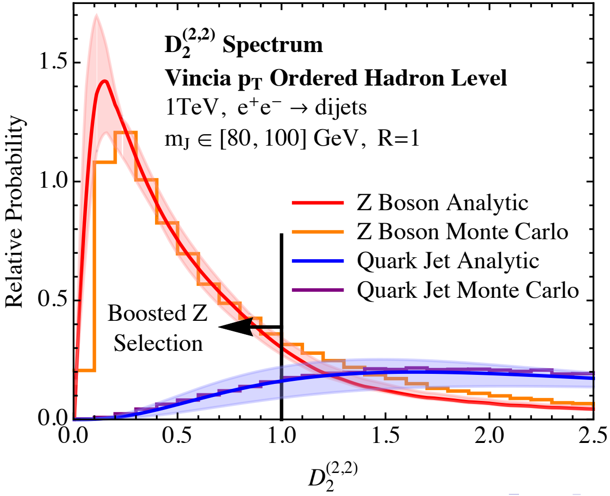
Analytic Boosted Boson Discrimination

- Analytic predictions for both signal and background allows for analytic boosted boson discrimination.



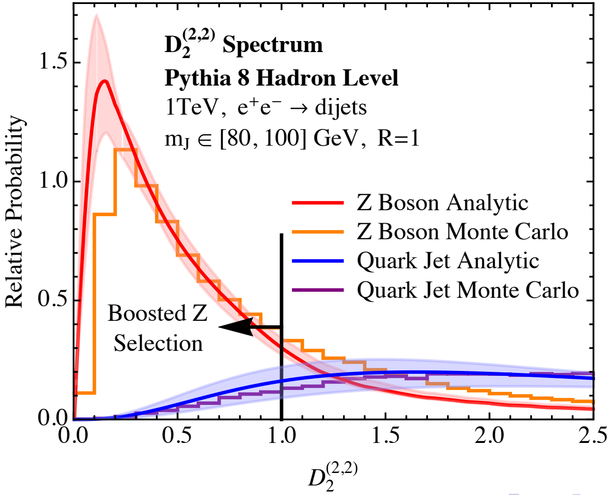
Analytic Boosted Boson Discrimination

- Region relevant for discrimination highly sensitive to non-perturbative physics. Excellent description with single parameter shape function!



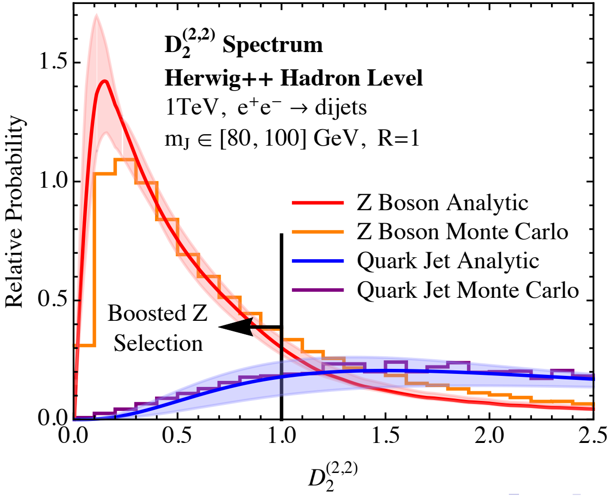
Analytic Boosted Boson Discrimination

- Region relevant for discrimination highly sensitive to non-perturbative physics. Excellent description with single parameter shape function!



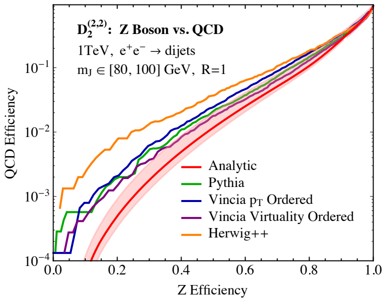
Analytic Boosted Boson Discrimination

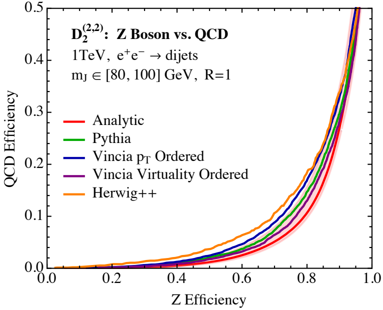
- Region relevant for discrimination highly sensitive to non-perturbative physics. Excellent description with single parameter shape function!



Analytic Boosted Boson Discrimination

- Analytic Signal Efficiency vs. Background Efficiency (ROC) curves

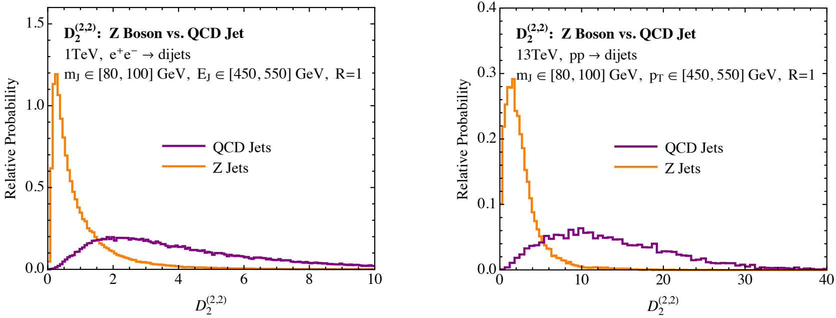




- Slightly optimistic prediction driven by signal calculation. Improveable with higher order resummation.

Future Directions

- Approach generalizes to other substructure observables of interest.
- Factorization theorem extends to pp. Here separation of perturbative and non-perturbative physics plays an even more crucial role.



Conclusions

- First rigorous all orders factorization theorem for a two prong observable valid for both signal and background (QCD) distributions.
- Separation of perturbative and non-perturbative physics. Field theoretic definition of non-perturbative parameters allows one to prove relations and maintain predictivity.
- Monte Carlo generators are not perfect. Analytic calculations essential to understand their abilities.

Thanks!