Utilizing Color Flow for Discovery

BNL Forum Parallel Talk

May 2 2013

David Curtin
Yang Institute for Theoretical Physics
Stony Brook

Based on

1210.5523, **work in progress** (DC, Rouven Essig, Brian Shuve)

Introduction / Motivation

- Lots of progress in jet substructure over the last few years.
- 'Historical' focus has been on identifying boosted EW mass scale objects (top, W, Z) as 'detector objects' at the LHC (much like b-jets that can be tagged) and removing Pile-Up.
  - This lead to many tools being successfully adopted by experimentalists. (HEPTopTagger, BDRS, N-subjettiness, jet grooming, ...)

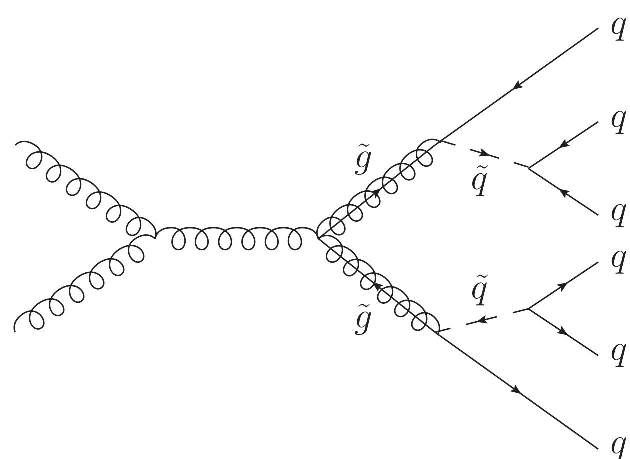
    → Tend to be based on '**hard substructure**' (splittings, mass drops inside a fat jet)

See e.g. BOOST proceedings

- Recently, a lot of '**soft substructure**' being developed by theorists that probes the shape of a jet's radiation field (girth, R-cores, color-flow,...)
  - Largely **not** yet experimentally verified/adopted.
  - **How useful is soft substructure at the LHC?**

RPV Gluinos

- In light of non-discovery of SUSY, RPV models are interesting since they lack the large MET signatures of RPC SUSY.
- A very difficult signature is gluino → 3 light quarks via an off-shell squark. 6j signal, two 3j resonances.
  → combinatorics!
  → QCD multijet background!



- Has been searched for by Tevatron and CMS using 6j resolved search that attempts to reconstruct the resonance. (excludes 77 - 144 & 200 - 450 GeV)
  1105.2815
  1208.3084
  ATLAS did a 6j counting experiment that excludes < 650 GeV
  1210.4813

Boosted RPV Gluinos

- Would like to do a **boosted gluino search** where each gluino forms a single fat jet. This eliminates combinatorics background and **reliably** reconstructs the resonance.
  - Looking for two fat jets with similar masses, each containing three similarly hard subjets could have raised Tevatron limit from 140 GeV to ~250 GeV

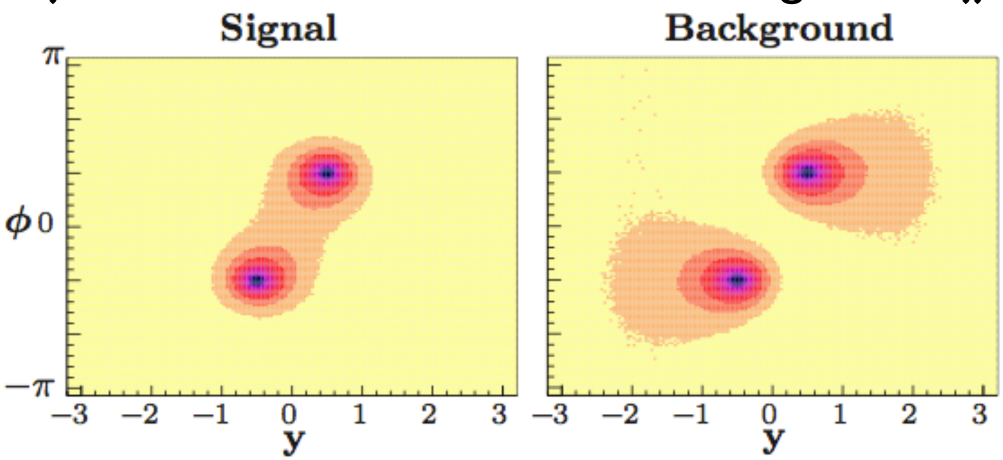
    Raklev, Salam, Wacker 2010

- We would like to revisit this issue at the LHC with new methods.
  - **Can boosted beat resolved/counting?**
  - **Can soft substructure help?**
    - **The off-shell squark means gluino forms R-hadron!**

This produces a composite **color singlet** with **strong production cross section** that decays into **three jets**. **Very unique signal!**

Color Flow

- If a color-singlet decays to two quarks, the resulting color-dipole will have a radiation pattern that is concentrated between the two jets, different from QCD jets which are beam-connected. E.g. . Z → jj

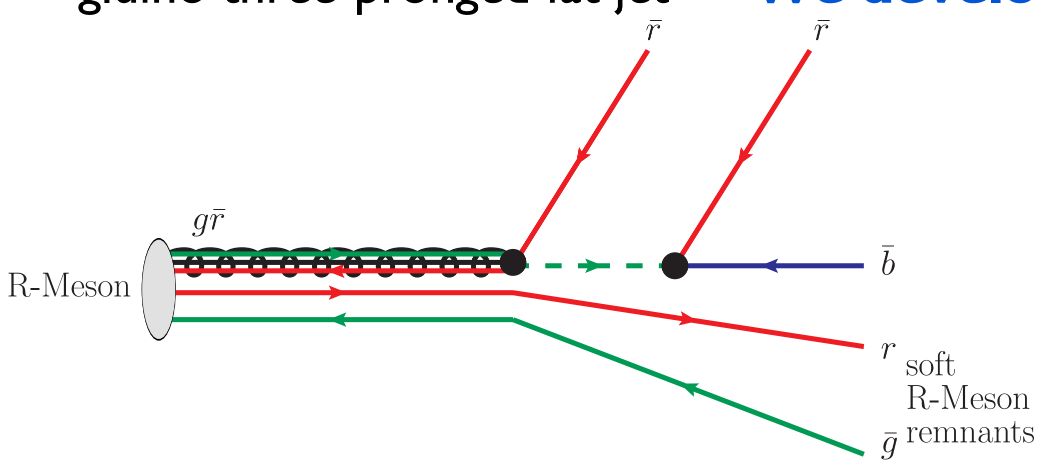


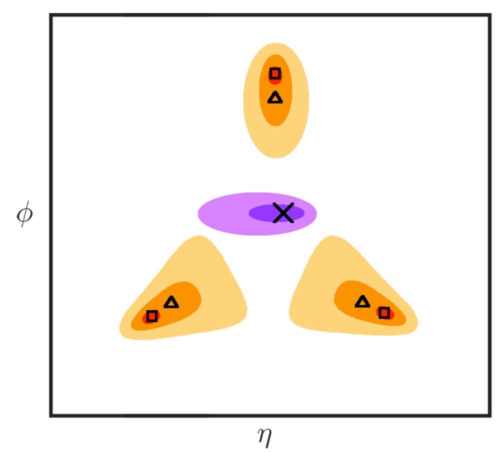
Can probe this with various variables, e.g. pull, dipolarity

$$\vec{t} = \sum_i \frac{p_T^i |r_i|}{p_T^{jet}} \vec{r}_i$$

1001.5027 (Gallicchio, Schwartz), 1102.1012 (Hook, Jankowiak, Wacker), ...

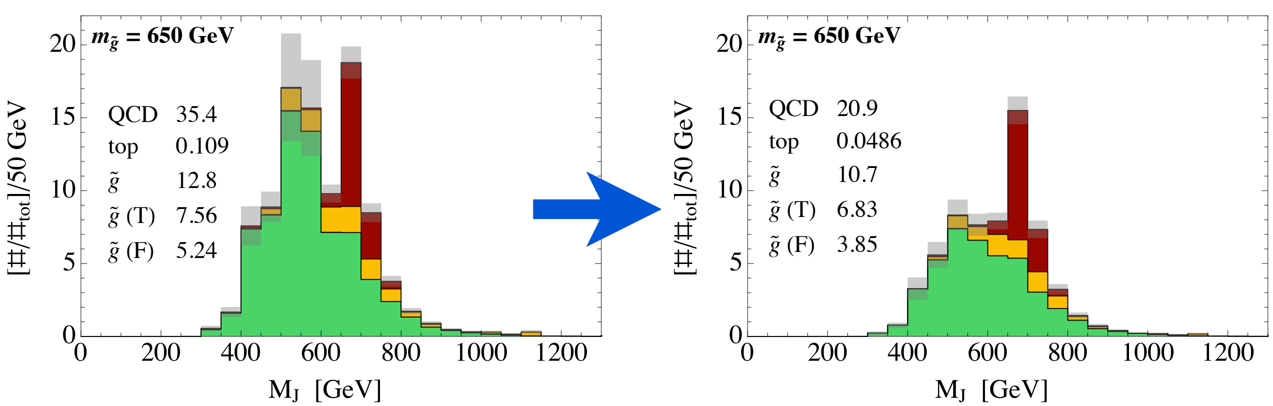
- The hard AND soft decay products of an **R-hadron** have to form color singlets as well. Naively, would expect this to yield a **unique radiation pattern** in the gluino three-pronged fat jet → **We develop variables to probe this.**





Boosted Search: Heavy Gluino

- Generate QCD BG in Sherpa, Gluino Signal in Pythia 8
- Apply kinematic cuts (two high-pT fat jets with similar mass and three similarly hard subjets)
- Small signal: boosted gluino fraction ~ few %. O(10%) of that SIGNAL survives our cuts to give S/B ~ 1.
- Cut on Color Flow (Axis Contraction) cleans up distribution.



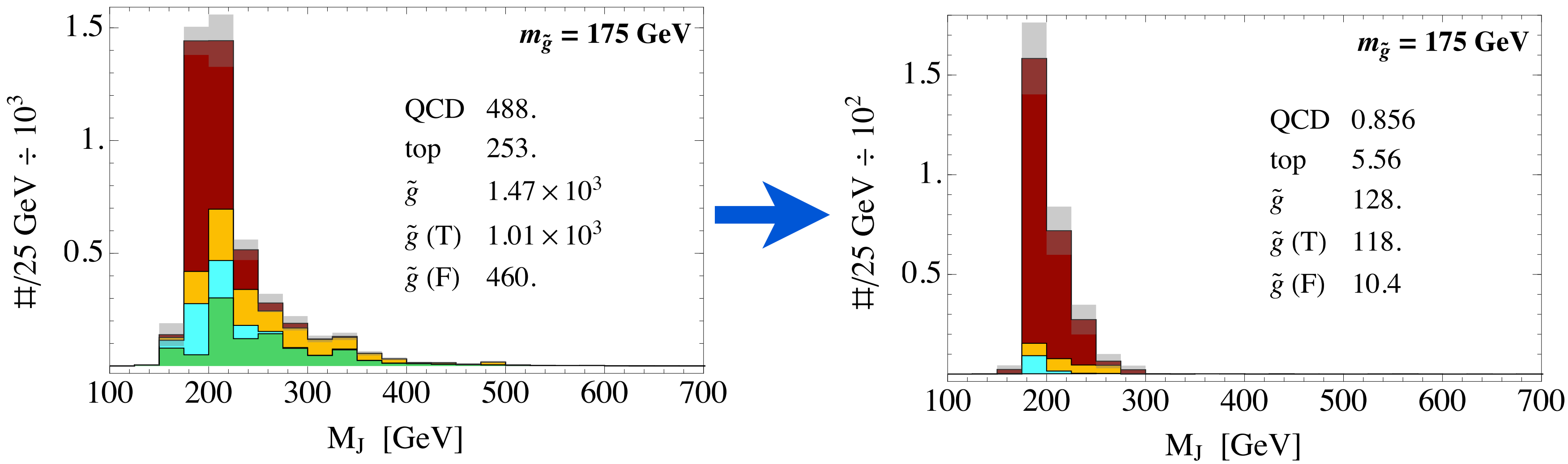
Could obtain bounds
650 GeV (LHC8 5/fb)
750 GeV (LHC8 20/fb)

Our boosted search has same reach as ATLAS counting experiment!

Boosted Search: Light Gluino

- Top-Mass Gluino demonstrates power of color-flow cuts: Large signal → can cut very hard to obtain high purity

- Aggressive cut on Color Flow (Radial Pull): S/B ~ 3 → 100!

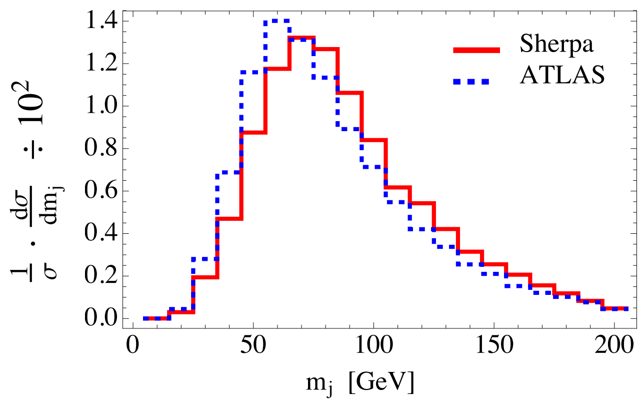


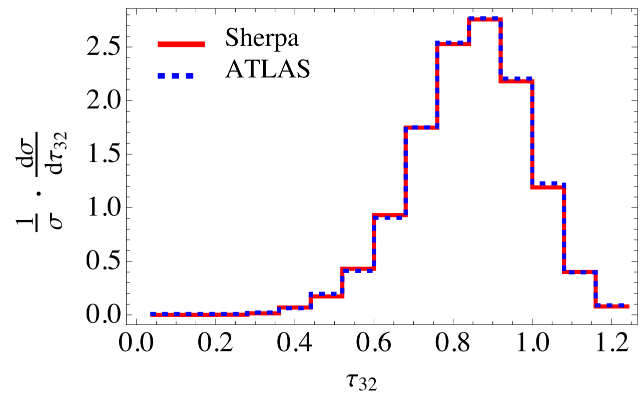
- **Demonstrates scenario where color flow is very powerful signal discriminator!**

(We also checked lots of other substructure variables like girth, planar flow, .. and none of them were any use here.)

"Aside": MC Validation

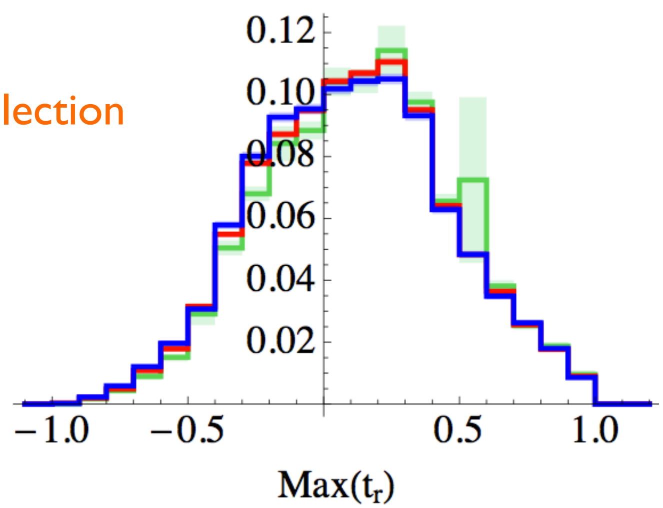
- QCD background challenging to model. Normalize to DATA (35/pb ATLAS measurements). Excellent shape agreement.



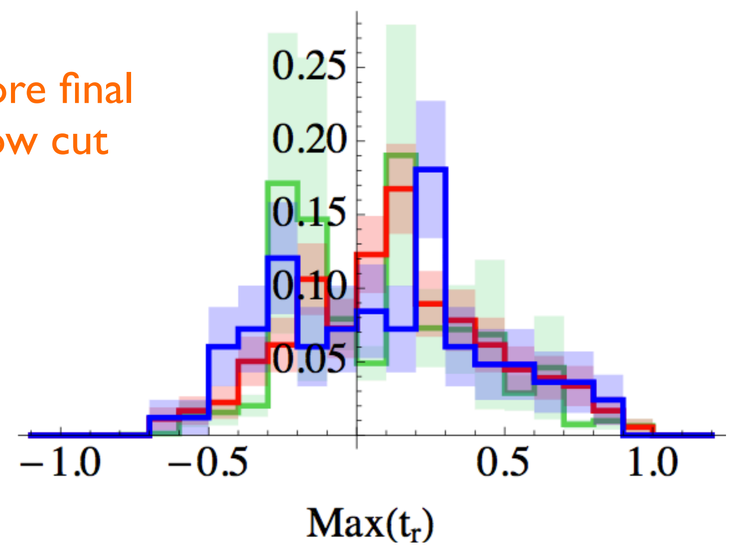


- Compared distributions in Sherpa to POWHeg + Pythia6.4, POWHeg + Pythia8. Shape agreement is generally good, but some important deviations in cut efficiencies & tails of distributions for color flow.

After preselection



Just before final color-flow cut



What have we learned so far?

- A boosted gluino search could do as well as a counting experiment, but with the more convincing mass peak!
- Color flow cleans up heavy gluino case and could have lead to spectacular results for top-mass gluinos.
- **Can color flow (and other soft substructure variables) be helpful in a less `extreme' scenario?**
- It would be great to have a `Killer App' for soft substructure that can serve as a guide-post to motivate all the extra work they still need to become experimentally viable.
    - What cannot be done without soft substructure?

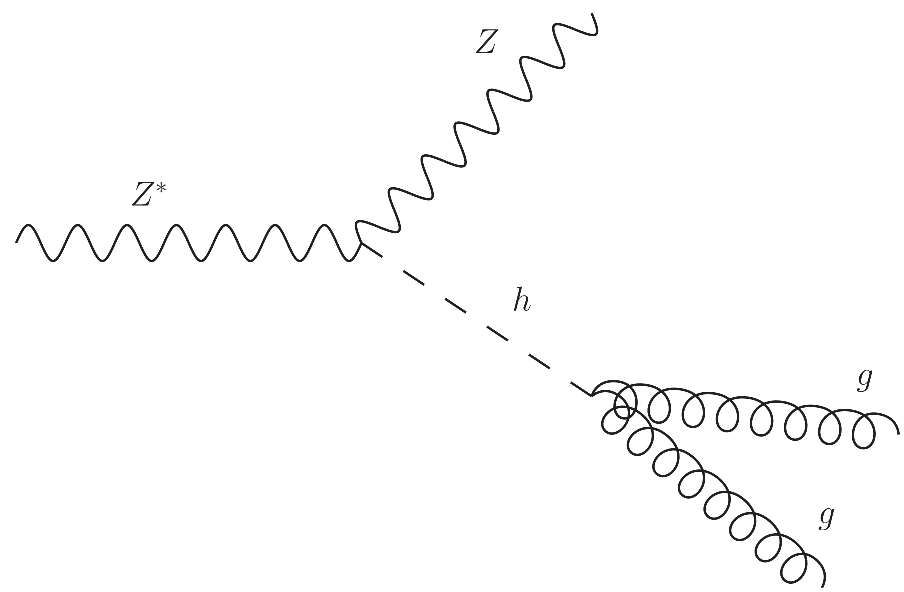
What have we learned so far?

- A boosted gluino search could do as well as a counting experiment, but with the more convincing mass peak!
- Color flow cleans up heavy gluino case and could have lead to spectacular results for top-mass gluinos.
- **Can color flow (and other soft substructure variables) be helpful in a less `extreme' scenario?**
- It would be great to have a `Killer App' for soft substructure that can serve as a guide-post to motivate all the extra work they still need to become experimentally viable.
  - What cannot be done without soft substructure?

How about h→gg?

$h \to gg$: Soft Substructure Show-Pony?

- Imagine boosted higgs analysis (like BDRS) at high-lumi LHC
    - replace b-tags by some multivariate soft-substructure tagger to lock onto di-gluon color singlet from higgs decay
- Many motivations to measure this (SM closure test, $h \to gg$ vs $h \to bb$ ratio)
- Di-gluon color-singlet is a very special state, color flow should work very well.
- Possible problems: **Wjj BG** is huge, and Wgg has a color-singlet component that represents irreducible background.



Also: **Pile-Up!**

$h \to gg$ Preliminary Study

- Toy-Analysis in Madgraph + Pythia 8 for LHC14.
- Signal: $Wh \to l\nu\, gg$ 41 fb
  Main BG: $Wjj \to l\, \nu\, jj$ ~10 nb

S/B ~ 1/250k.

- To identify boosted $h \to gg$, some obvious kinematic cuts:
  - ✦ Tag on lepton from very hard W
    $p_T^l > 20$ GeV, MET > 30, $M_T^W < 90$ GeV, $p_T^W > 200$ GeV
  - ✦ Two-pronged fat jet (1.2) with pT > 200 GeV. Require hardest two thin jets to lie in this fat jet.

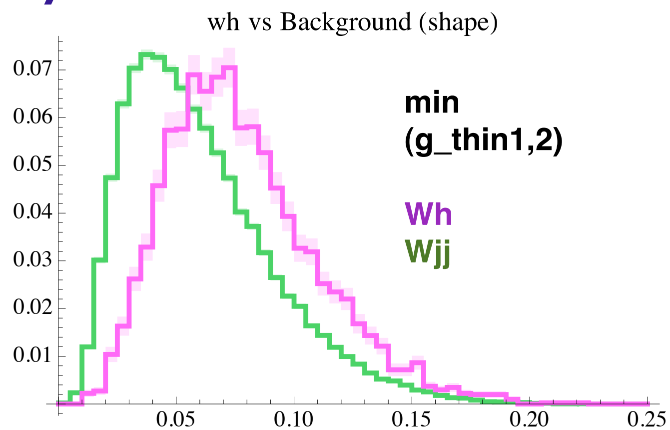
➡ In m_h-window, just from these kinematic cuts:
$S/B \sim 1/4000$
$S/\sqrt{B} \sim 0.4$ with 3000/fb

Need to do better!

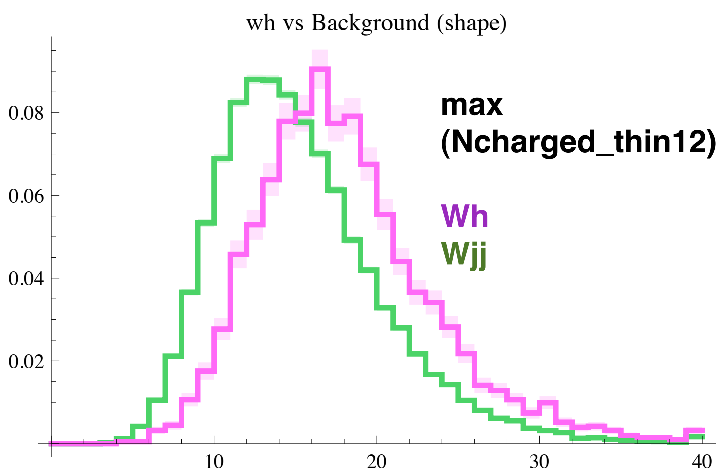
h→gg Preliminary Study

Can soft substructure get us up to 3σ sensitivity?

\`fuzzyness‘ (girth, R-cores, ..)

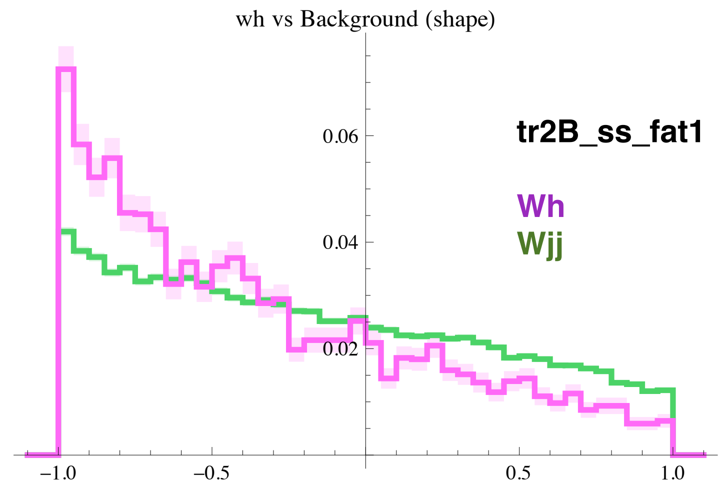


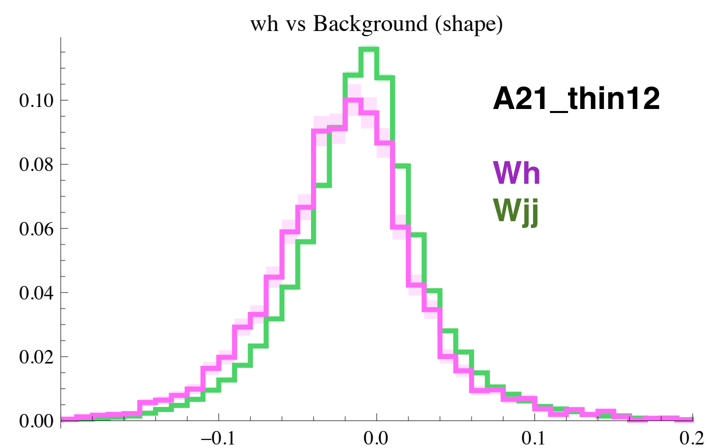
track count



Color Flow

(dipolarity, pull, axis contraction, color connectivity...)

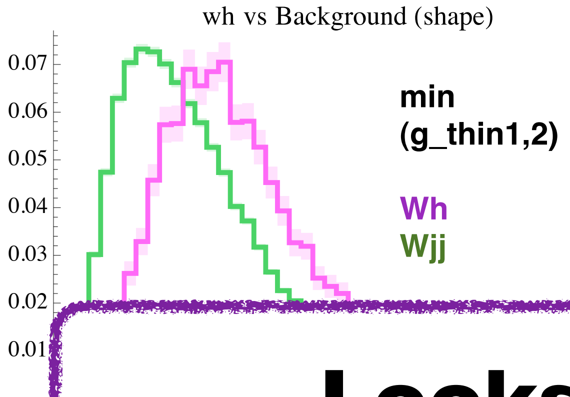




h→gg Preliminary Study

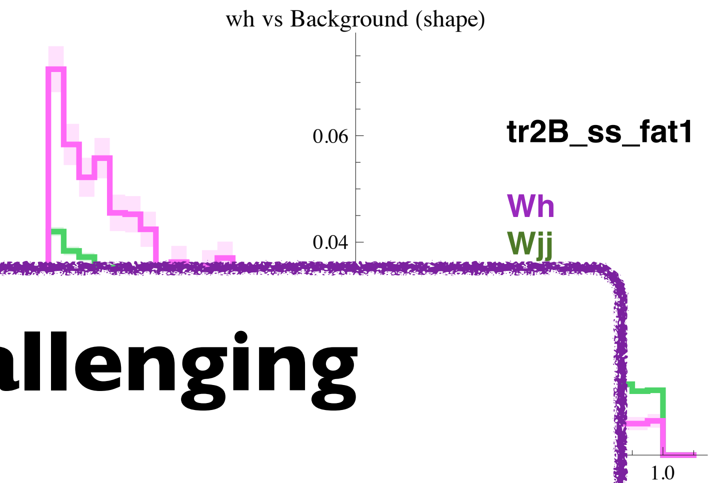
Can soft substructure get us up to 3σ sensitivity?

`fuzzyness‘ (girth, R-cores, ..)



Color Flow

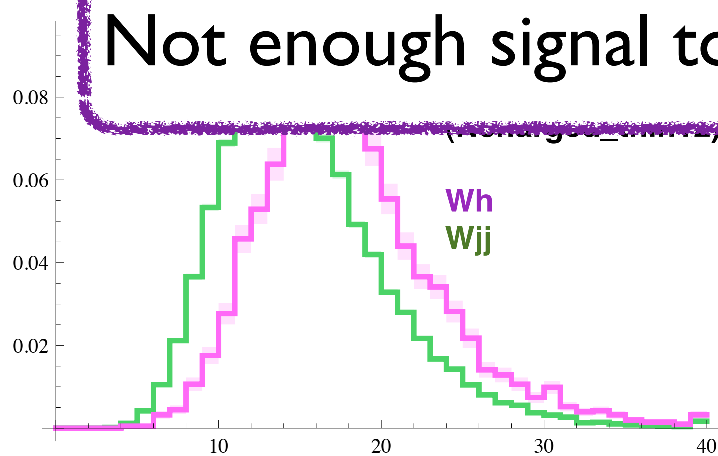
(dipolarity, pull, axis contraction, color connectivity...)

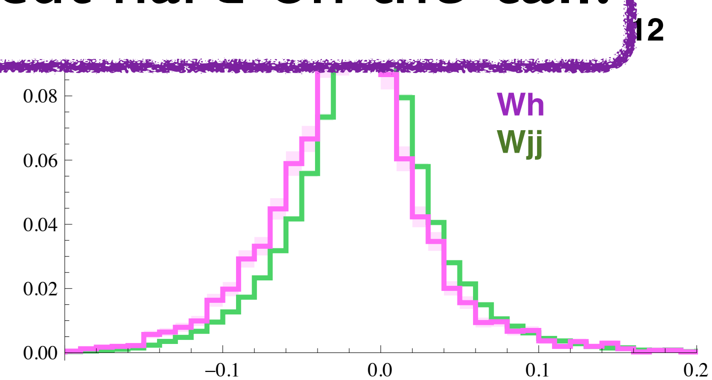


Looks very challenging

Not enough signal to simply cut hard on the tail.

track





Interpreting this Preliminary Result

- $h \to gg$ investigation is ongoing!
  - Should really separate Wgg, Wqq, Wqg. (Maybe the irriducible BG is just too large?)
  - Could multi-variate techniques help?
  - Pythia is a little pessimistic. Do things change in e.g. Sherpa? If yes, which is correct?
- Depending on the outcome, what's the story?
  - If $h \to gg$ can be measured: obviously amazing. Now theorists and experimentalists have a guiding goal to develop these methods.
  - If $h \to gg$ can NOT be measured:
    - Are there other well-motivated BSM scenarios where soft substructure is useful? Maybe a larger signal than $h \to gg$?

Interpreting this Preliminary Result

- $h \to gg$ investigation is ongoing!
  - Should really separate Wgg, Wqq, Wqg. (Maybe the irriducible BG is just too large?)
  - Could multi-variate techniques help?
  - Pythia is a little pessimistic. Do things change in e.g. Sherpa? If yes, which is correct?
- Depending on the outcome, what's the story?
  - If $h \to gg$ can be measured: obviously amazing. Now theorists and experimentalists have a guiding goal to develop these methods.
  - If $h \to gg$ can NOT be measured:
    - Are there other well-motivated BSM scenarios where soft substructure is useful? Maybe a larger signal than $h \to gg$?

Maybe it would be time to ask some more \`global' questions..

Disentangling Kinematics

- How much information is really contained in soft substructure?
- Would like to disentangle kinematic information!
- One possible approach:
  - Compare e.g. $Z \to jj$ to QCD dijets
  - **Artificially** adjust parton-level kinematics to be identical!
  - Can now ask very detailed questions:
    - ➡ How much S/B separation does soft substructure provide?
    - ➡ Assess volatility of variables (variation from shower)
    - ➡ Assess correllations amongst variables. Obtain minimal set to extract all information

Disentangling Kinematics

- How much information is really contained in soft substructure?
- Would like to disentangle kinematic information!
- One possible approach:
  - Compare e.g. $Z \to jj$ to QCD dijets
  - **Artificially** adjust parton-level kinematics to be identical!
  - Can now ask very detailed questions:
    - ➡ How much S/B separation does soft substructure provide?
    - ➡ Assess volatility of variables (variation from shower)
    - ➡ Assess correllations amongst variables. Obtain minimal set to extract all information

The outcome would be interesting no matter what the result!

Conclusions

- Jet substructure has evolved into a powerful tool at the LHC
- 'Soft' jet substructure is subtle and interesting
- Demonstrate its potential with RPV gluinos
  - Boosted search with resonance reconstruction and low S/B does as well as counting experiment without resonance information.
- In other applications, it is not clear how practically useful soft jet substructure really is
  - $h \to gg$ would be the killer app, but it looks very challenging
  - It may be time to answer some global questions:
    How much info in addition to kinematics?
    How volatile, how correlated? **In Progress...**