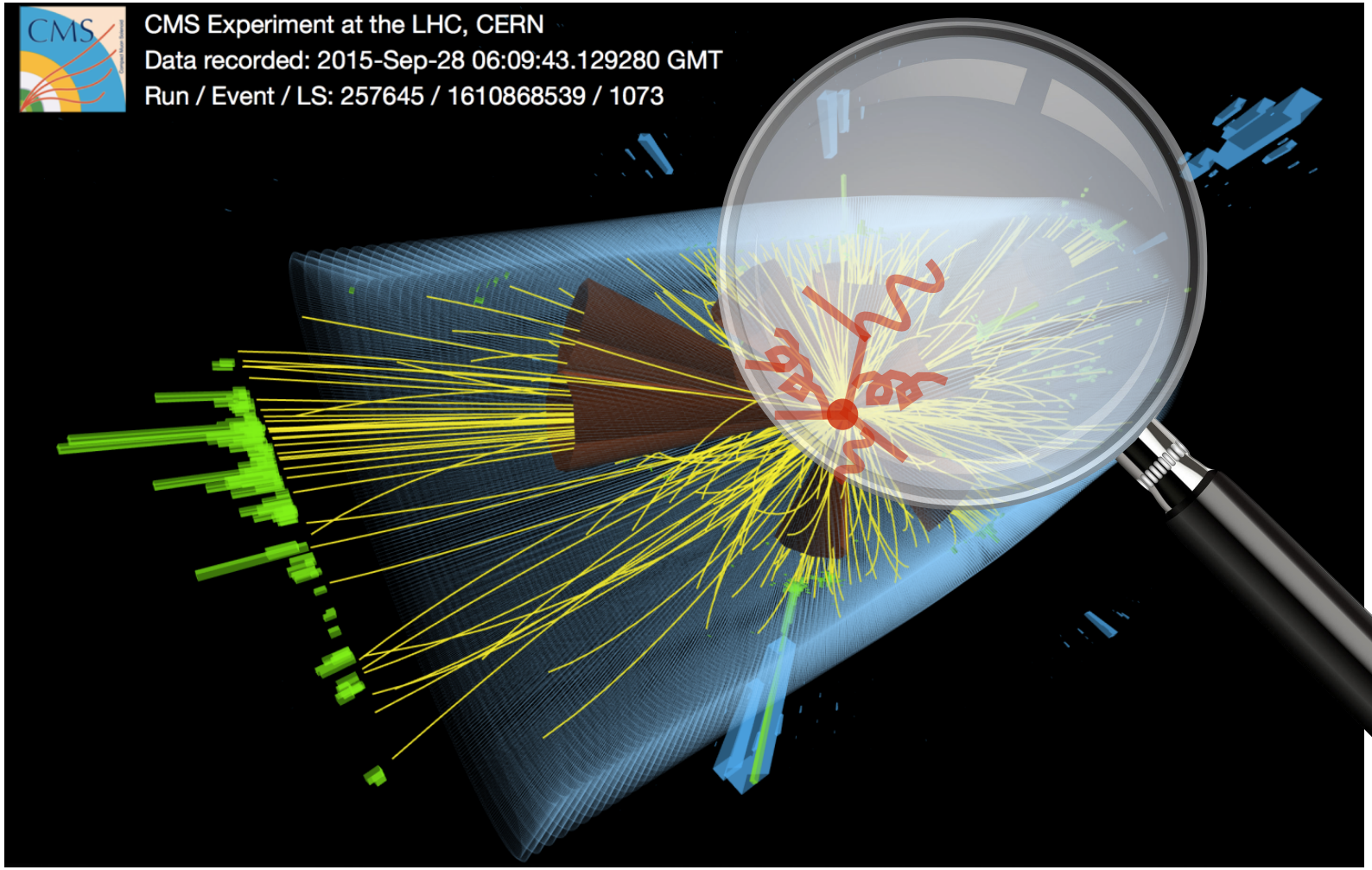


Resolving the Structure of New Physics: Tools for Discovery at the LHC

Justin Pilot, *UC Davis*
CMS Collaboration

Joint Experimental-Theoretical
Physics Seminar
Fermilab — 9 December 2016

The Search for New Physics

- The LHC and its experiments aim to uncover the fundamental nature of the universe
  - Are there additional symmetries and particles?
  - What is the nature of dark matter?
  - How did the universe evolve?
  - Are there extra dimensions of spacetime?
- You have seen many, many searches targeting those questions
  - Today's talk will focus more on the path to those results
  - How to use the detector information to improve our **discovery potential**

"…the horizon where the earth and the heavens meet…"

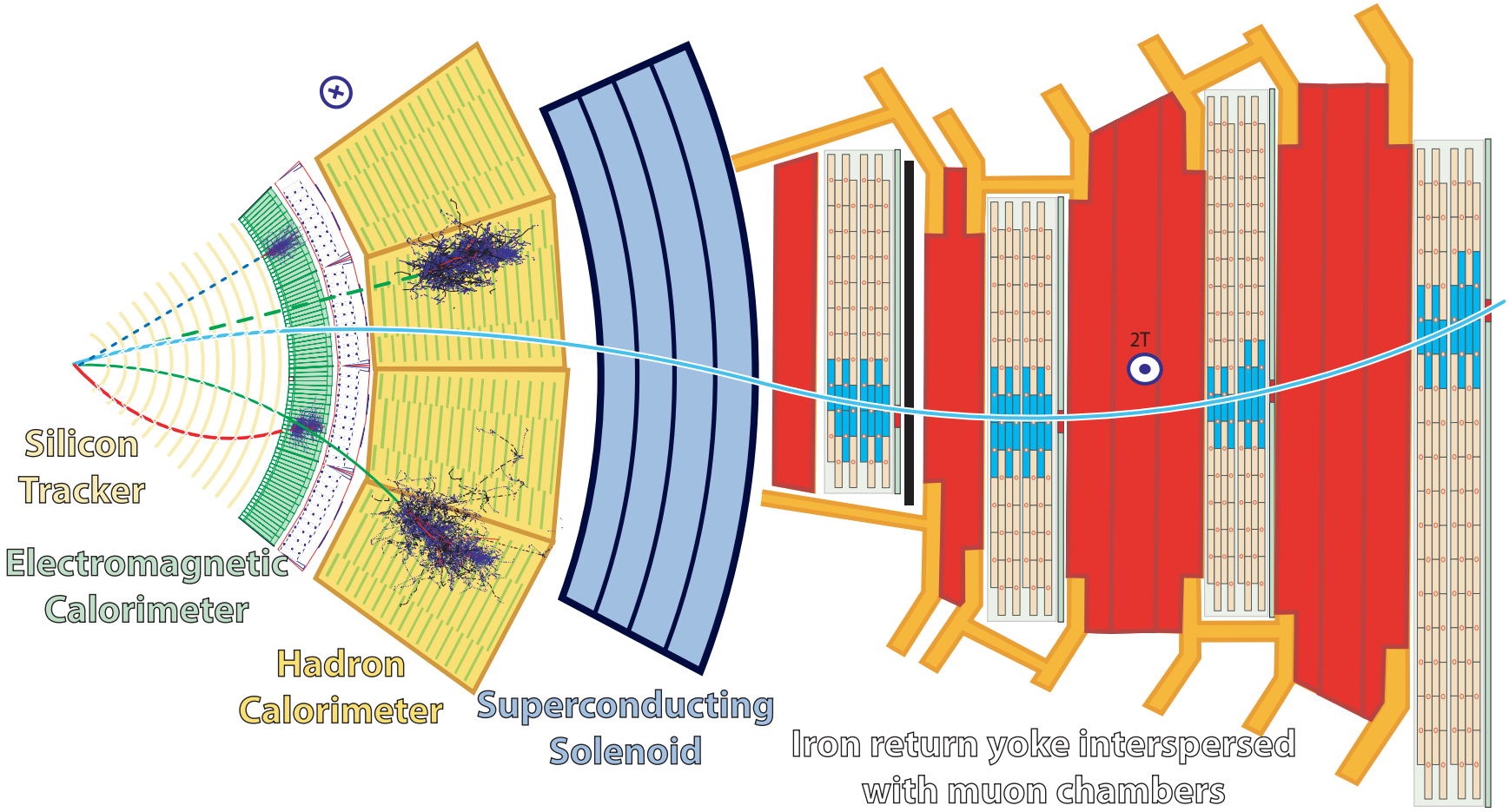
The Search for New Physics

- The LHC and its experiments aim to uncover the fundamental nature of the universe
  - Are there additional symmetries and particles?
  - What is the nature of dark matter?
  - How did the universe evolve?
  - Are there extra dimensions of spacetime?

- You have seen many, many searches targeting those questions
  - Today's talk will focus more on the path to those results
  - How to use the detector information to improve our **discovery potential**

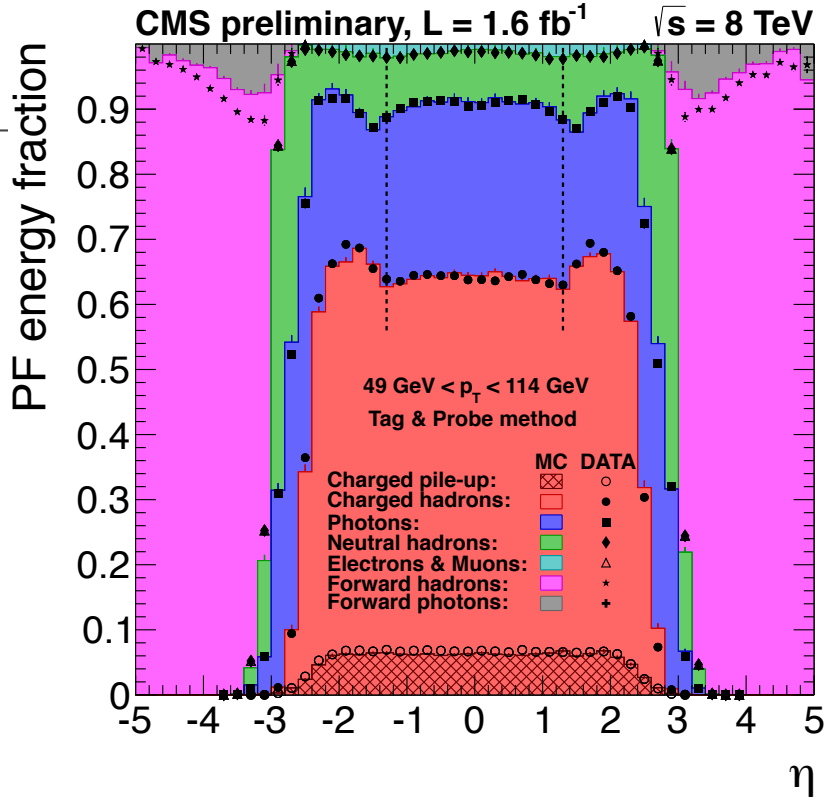
CMS Detector

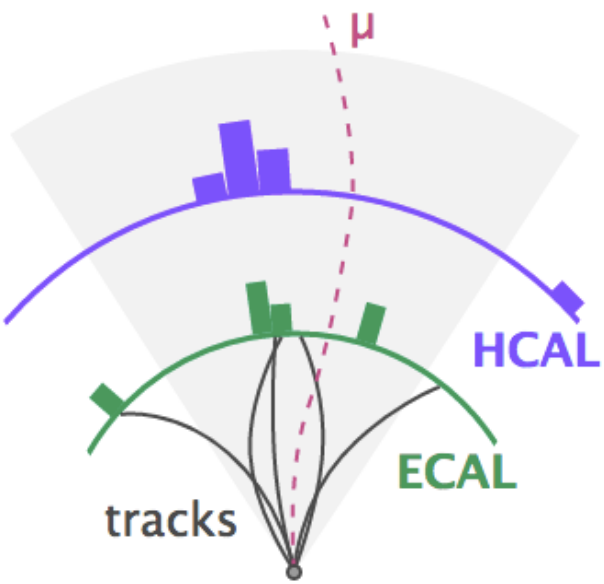
- Different particles; different detector signatures

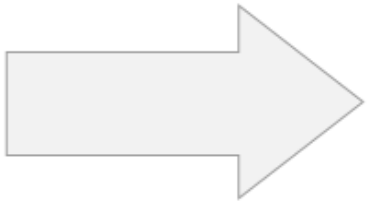


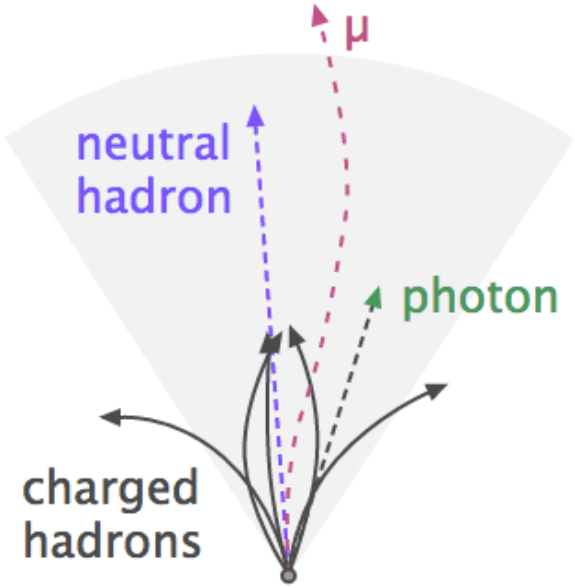
Event Reconstruction

- CMS uses the Particle Flow algorithm to reconstruct all particles produced in a collision event
  - Use information and signals from all detector components
- Particle information used to reconstruct physics objects







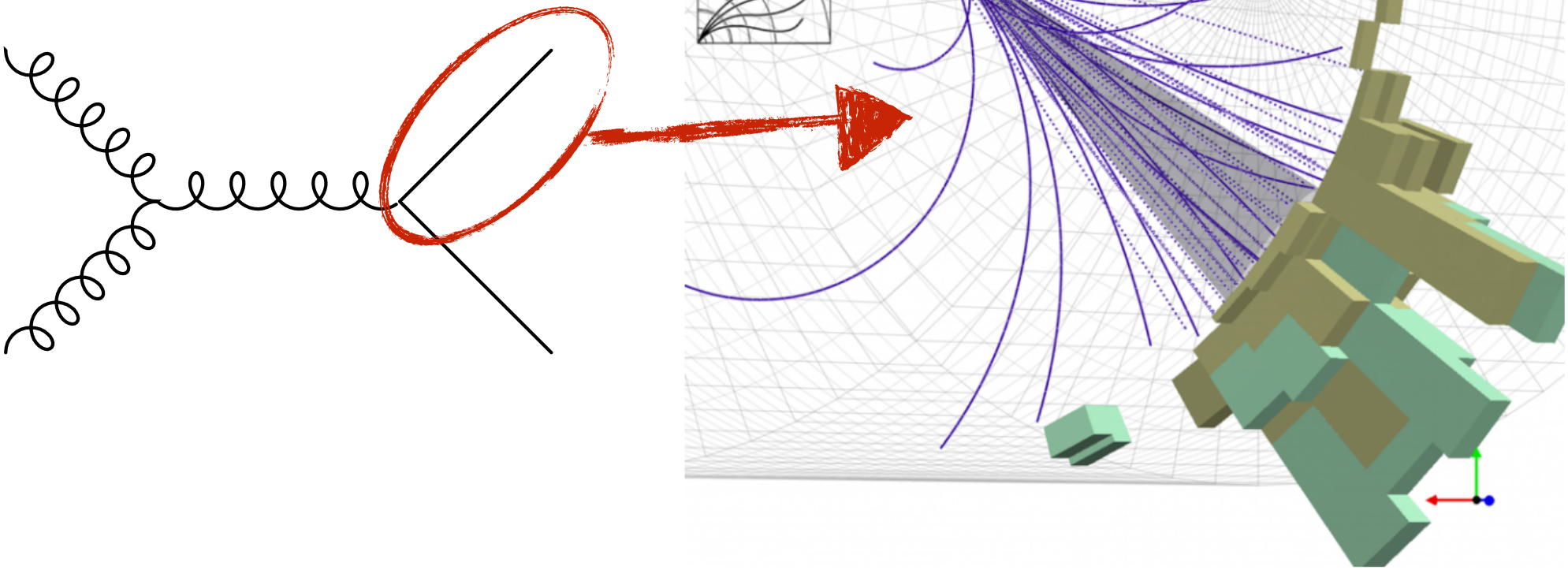


Detector level

Particle Flow

Event Reconstruction

- From this list of particles one can form **jets**, an object to reconstruct the shower of particles produced from a quark or gluon
- Each particle belonging to a jet is known as a **constituent**
  - Each has a 4-vector that can be used for further studies



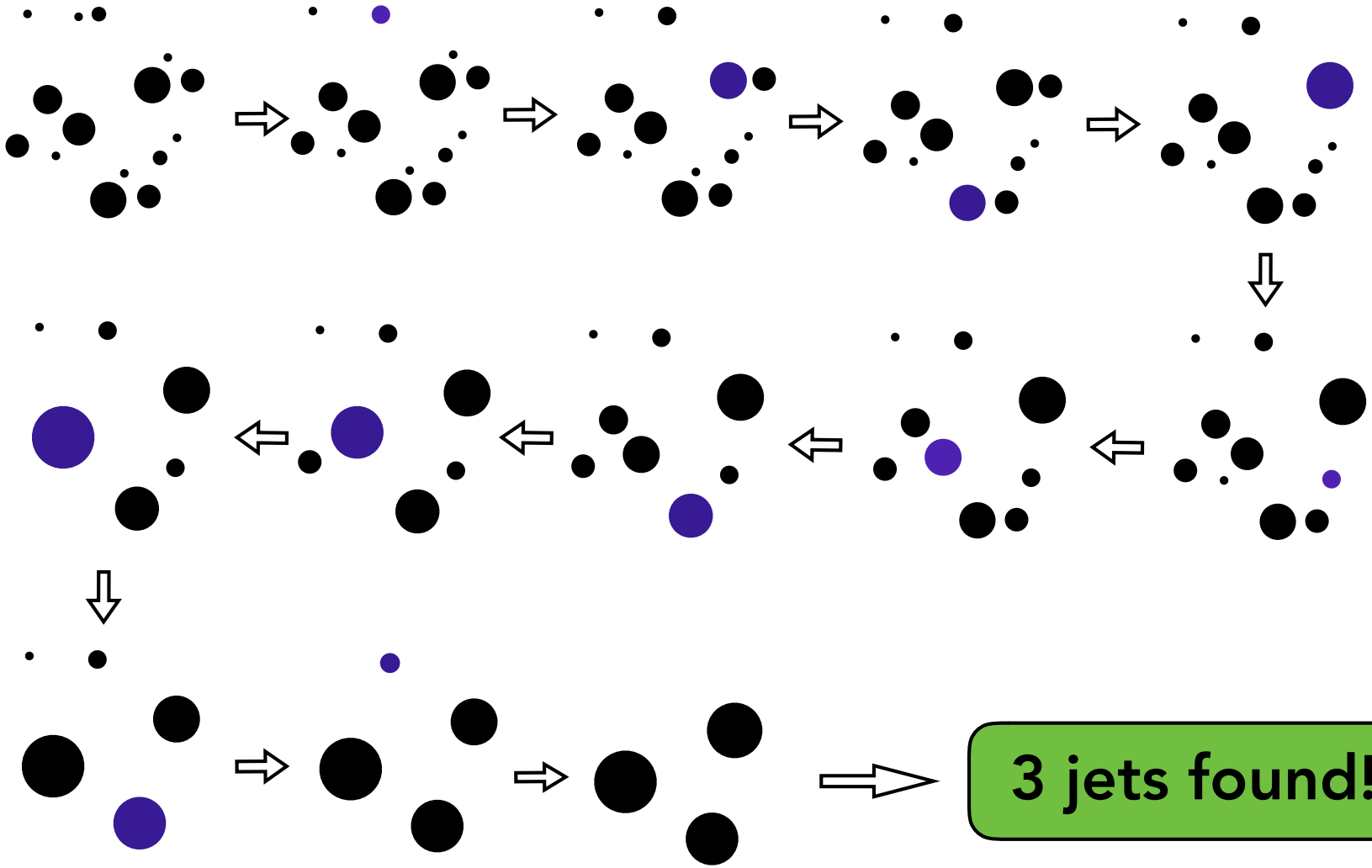
Jet Reconstruction

$$\beta = \begin{cases} -1, & \text{Anti-}k_T \\ 0, & \text{Cambridge-Aachen} \\ 1, & k_T \end{cases}$$

- Clustering algorithms used for a jet distance parameter *R* ("jet cone size")
  - If $d_{ij} < d_{ii}$, combine particles
  - If $d_{ii} < d_{ii}$, i is a jet

$$d_{ii} = p_{T,i}^{2\beta}$$

$$d_{ij} = \min(p_{T,i}^{2\beta}, p_{T,J}^{2\beta}) \frac{\Delta R_{ij}^2}{R^2}$$



3 jets found!

Jet Reconstruction

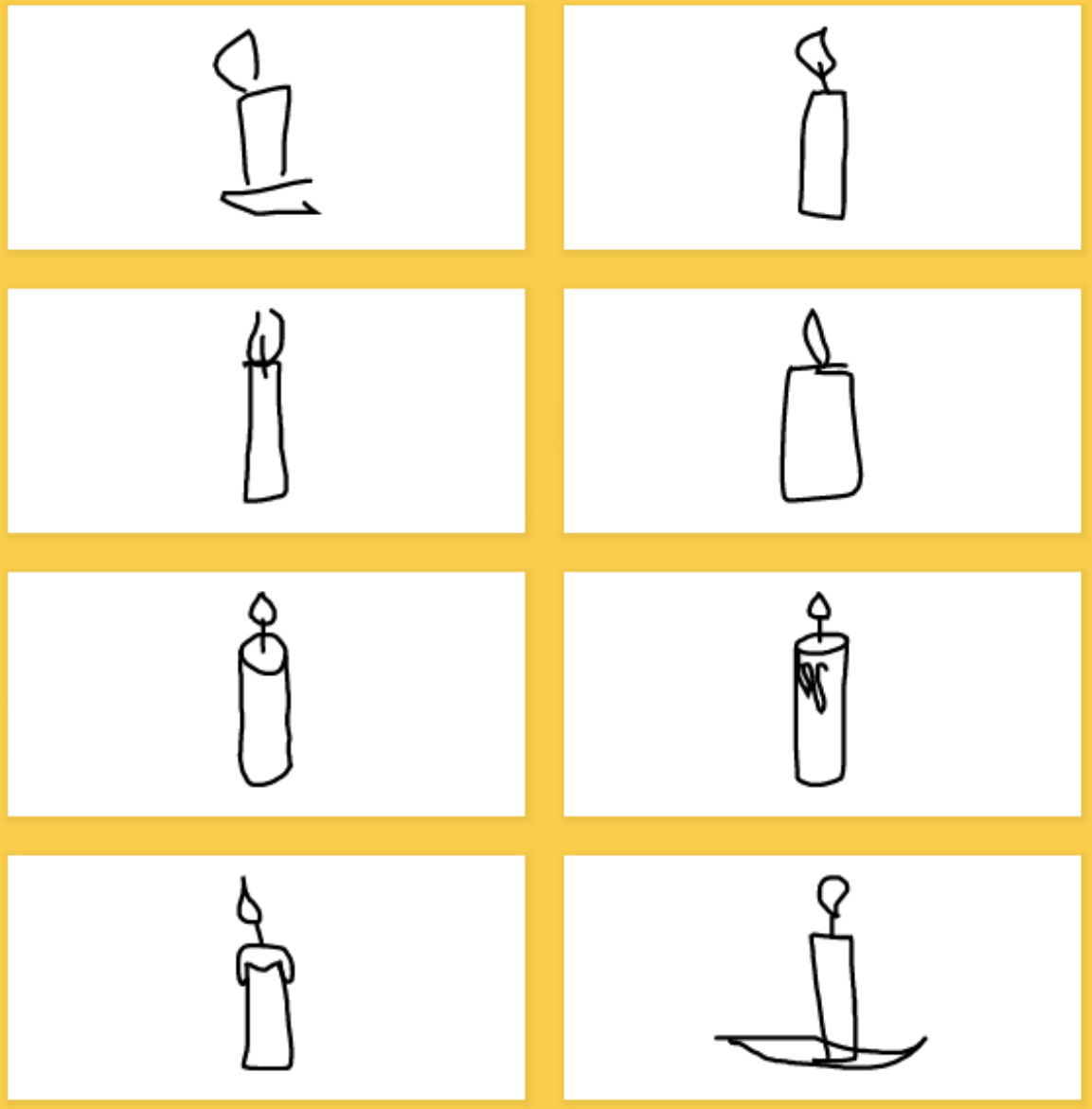
- Qualitatively different quarks/gluons produce different jet topologies
  - Different radiation patterns
  - Decay lifetimes
  - Can use topologies to discriminate
- Jets can also form from **hadronic decays of high-p_T heavy particles**
  - W/Z→qq, H→bb, t→Wb→qqb
- By looking at these patterns we can gain useful information to identify new physics signatures
  - **Jet substructure** techniques

An Analogy

- https://quickdraw.withgoogle.com/

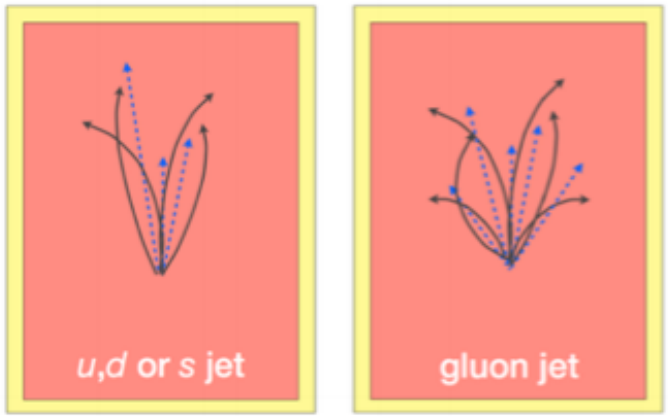
Easy or Hard?

How does it know what candle looks like?
It learned by looking at these examples drawn by other people.



Easy or Hard?


You were asked to draw cloud
It also thought your drawing looked like these:
Correct match
cloud
2ND closest match
bush
3RD closest match
sheep


u,d or s jet
gluon jet

Easy or Hard?

You were asked to draw **gluon jet**

You drew this, and the neural net didn't recognize it.

It thought your drawing looked more like these:

Closest match
snowflake

2ND closest match
garden

3RD closest match
campfire

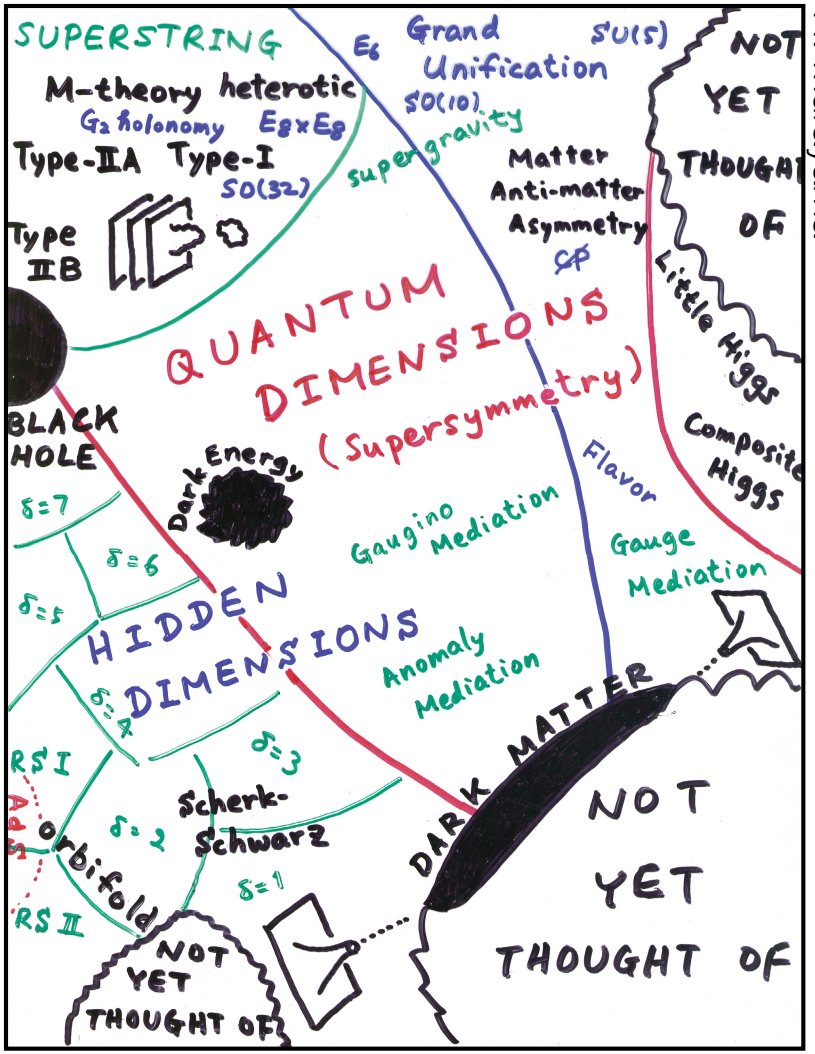
Discovery Tools

- Discovery of the Higgs leaves us with the Hierarchy Problem
  - Why is there such a large difference between the Higgs mass scale and the Planck scale
  - Some new physics must be present!

- Largest contribution to Higgs mass corrections is through a top quark loop
  - **New states related to the top quark are expected to be within reach**
    - Order ~few TeV
    - Models such as Composite Higgs, RS KK gluon, Extra dimensions, Topcolor Z', e.g.

- These new particles decay and produce **very high p_T SM particles**
  - Top quarks, W, Z, H bosons

- **Jet substructure signatures are discovery signatures**
  - Critical tools to maintain sensitivity in the high mass regimes

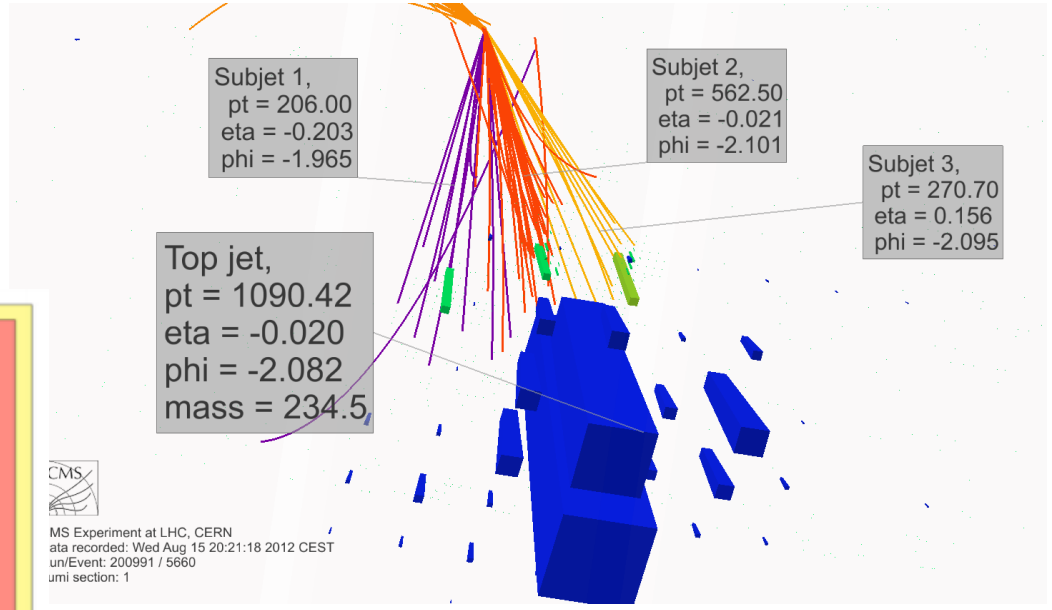


H. Murayama

Discovery Tools

- **Jet substructure signatures are discovery signatures**
  - Resolve individual decay products at high p_T
  - Mitigate degradation effects due to high pileup activity
  - Enhance search power for very high-mass particles

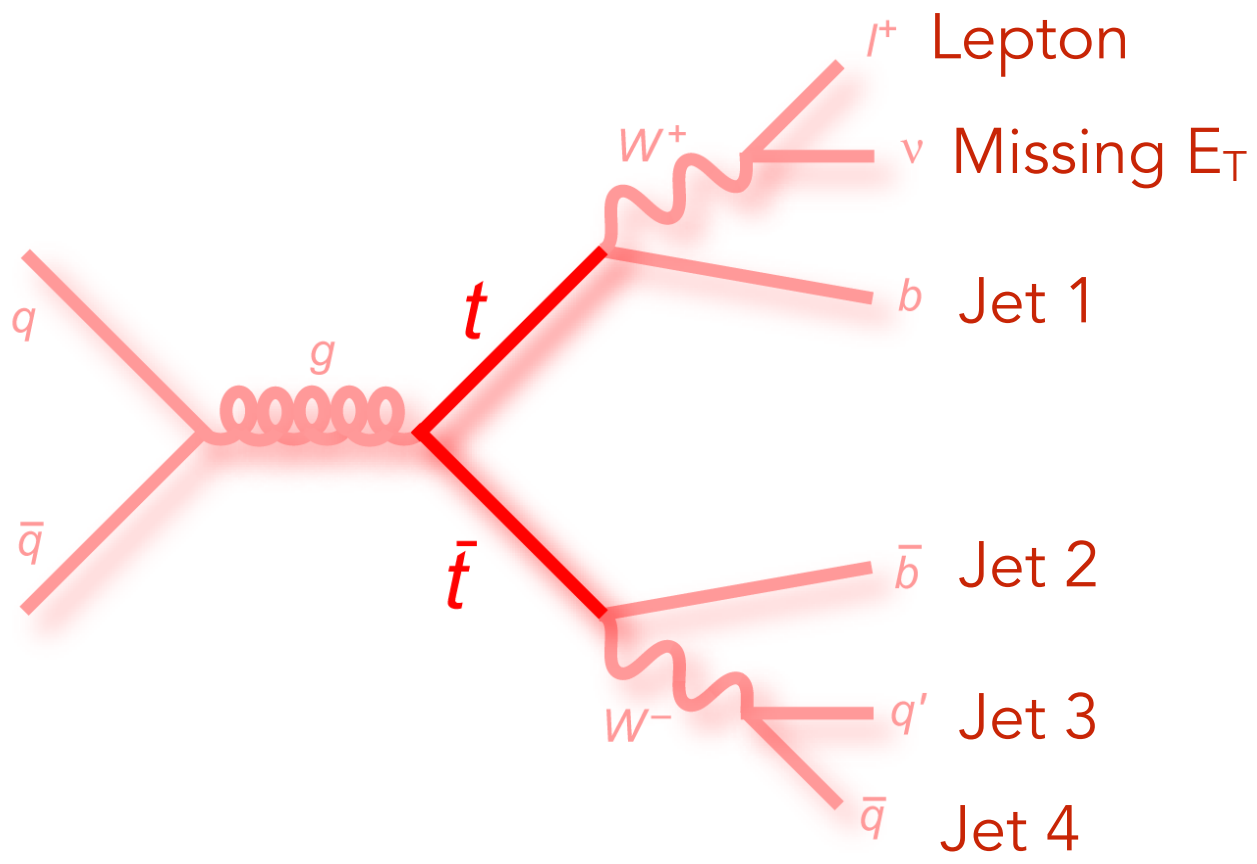
- Today I will discuss development of these algorithms and their use in CMS search analyses
  - Historical Perspective
  - Jet substructure algorithms
  - Analysis Highlights
  - Future Developments



Jet Substructure Development

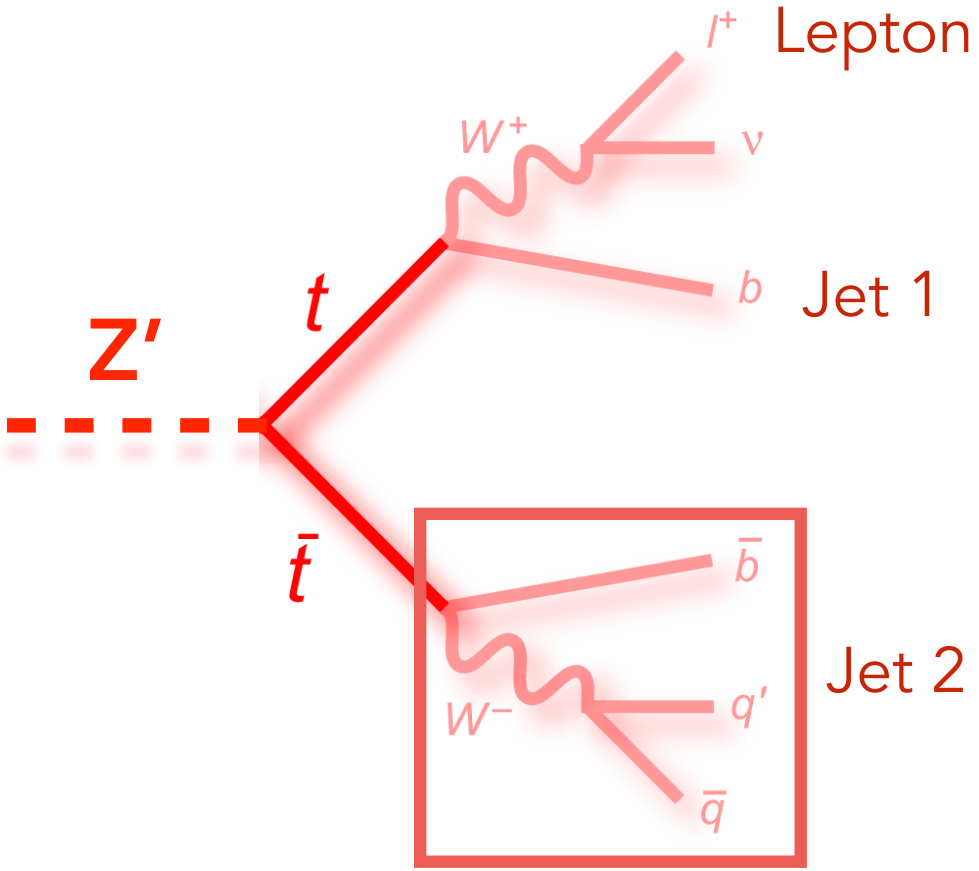
Historical Perspective

- Identifying SM ttbar events historically done by associating one object to each final state decay product
  - Combine objects to reconstruct each top quark
- Combinatorics can become unwieldy
  - 6+ jets in all-hadronic decay mode!

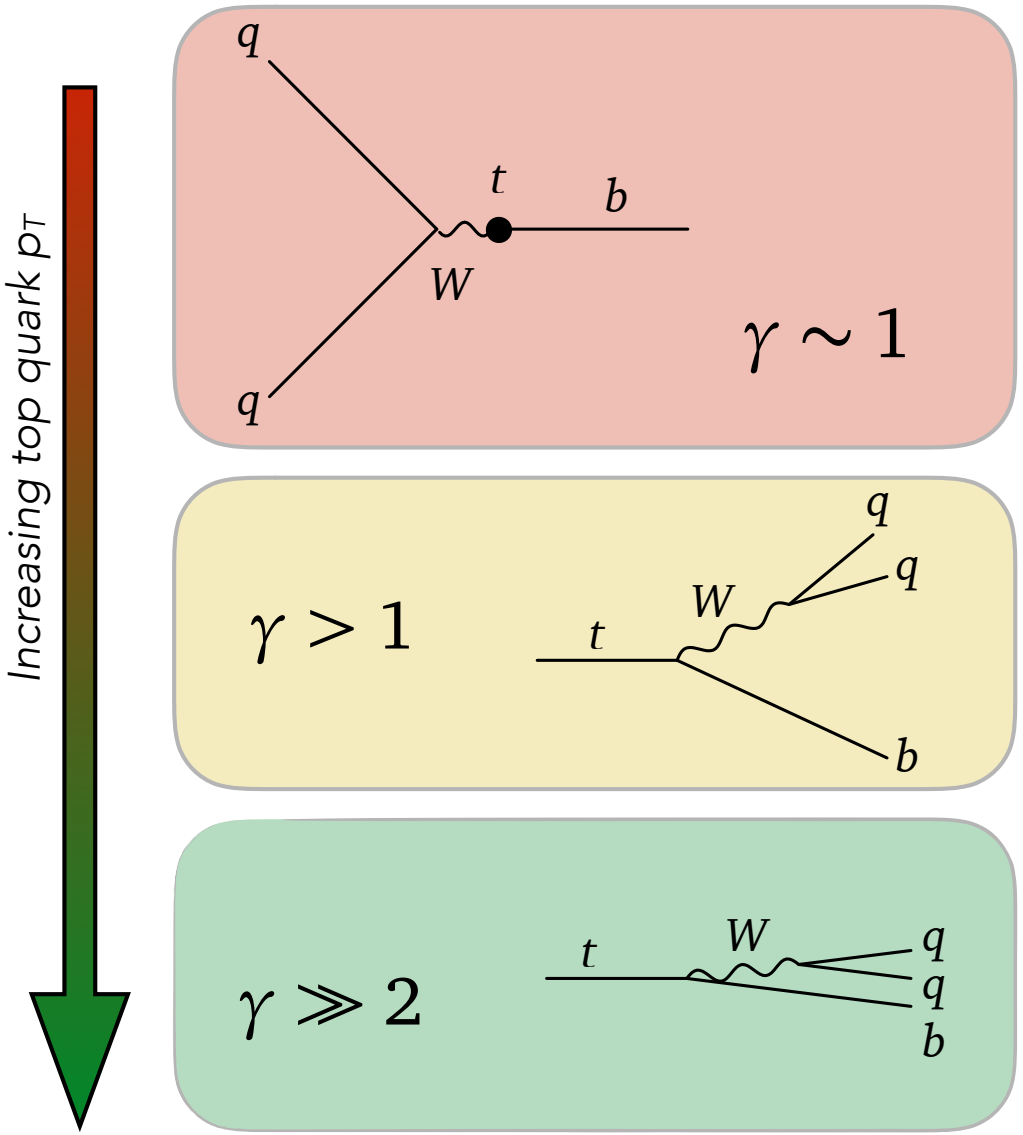


Historical Perspective

- If the top quarks are **boosted**, e.g. when coming from a new massive particle, what happens?
  - Hadronic decay products collimated → reconstructed in the same final-state object!

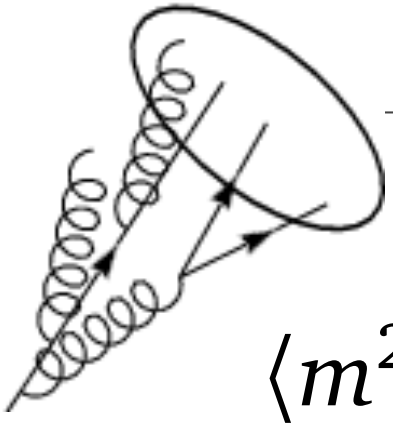


- Hadronic final states now become accessible with a dijet final state (in this case)



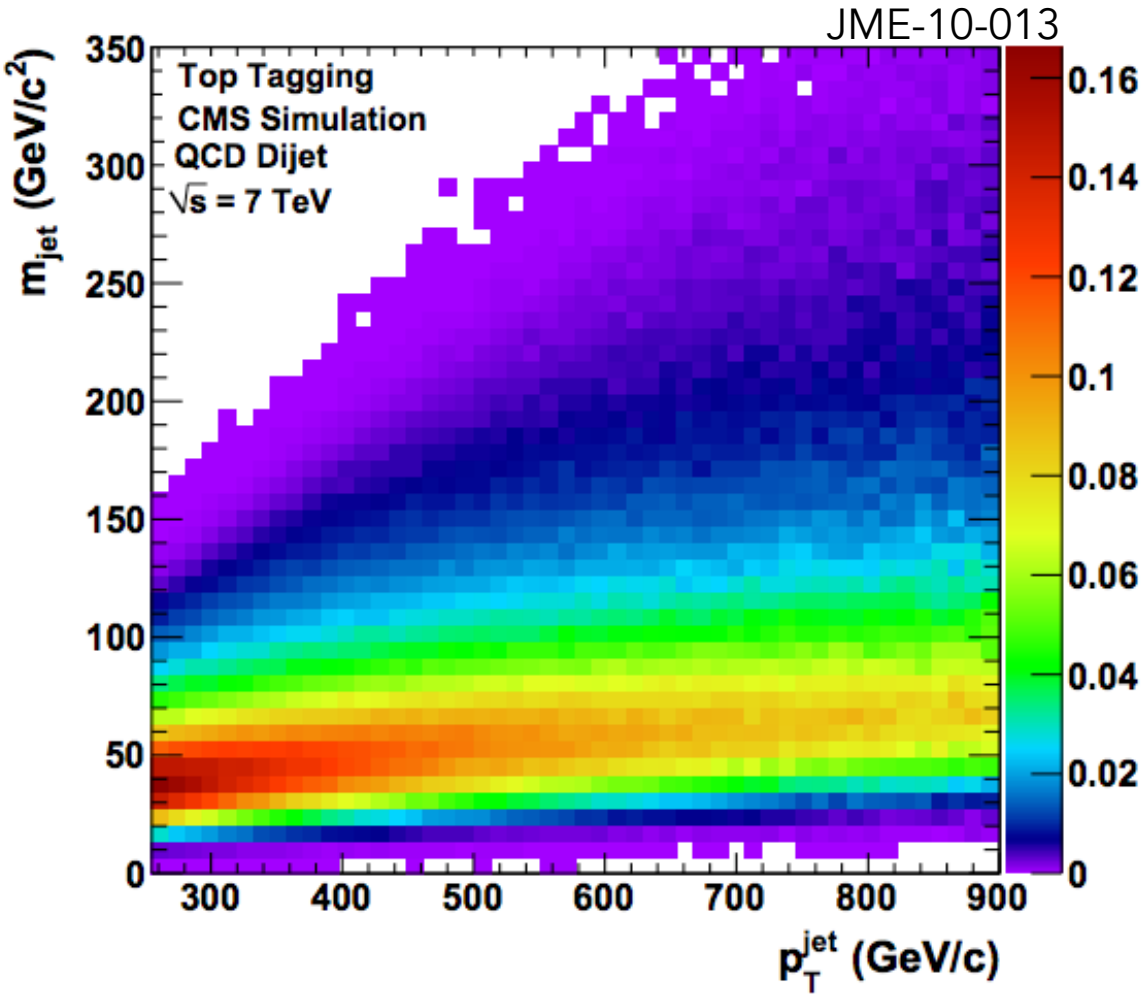
Jet Mass

- One important quantity to identify these jets is the jet mass
- Computed by adding up constituent particle 4-vectors and computing the mass
- For QCD jets, this depends on:
  - The momentum (p_T)
  - The size of the jet used (R)



$$\langle m^2 \rangle \sim \frac{C\alpha_S}{\pi} p_T^2 R^2$$

JME-10-013

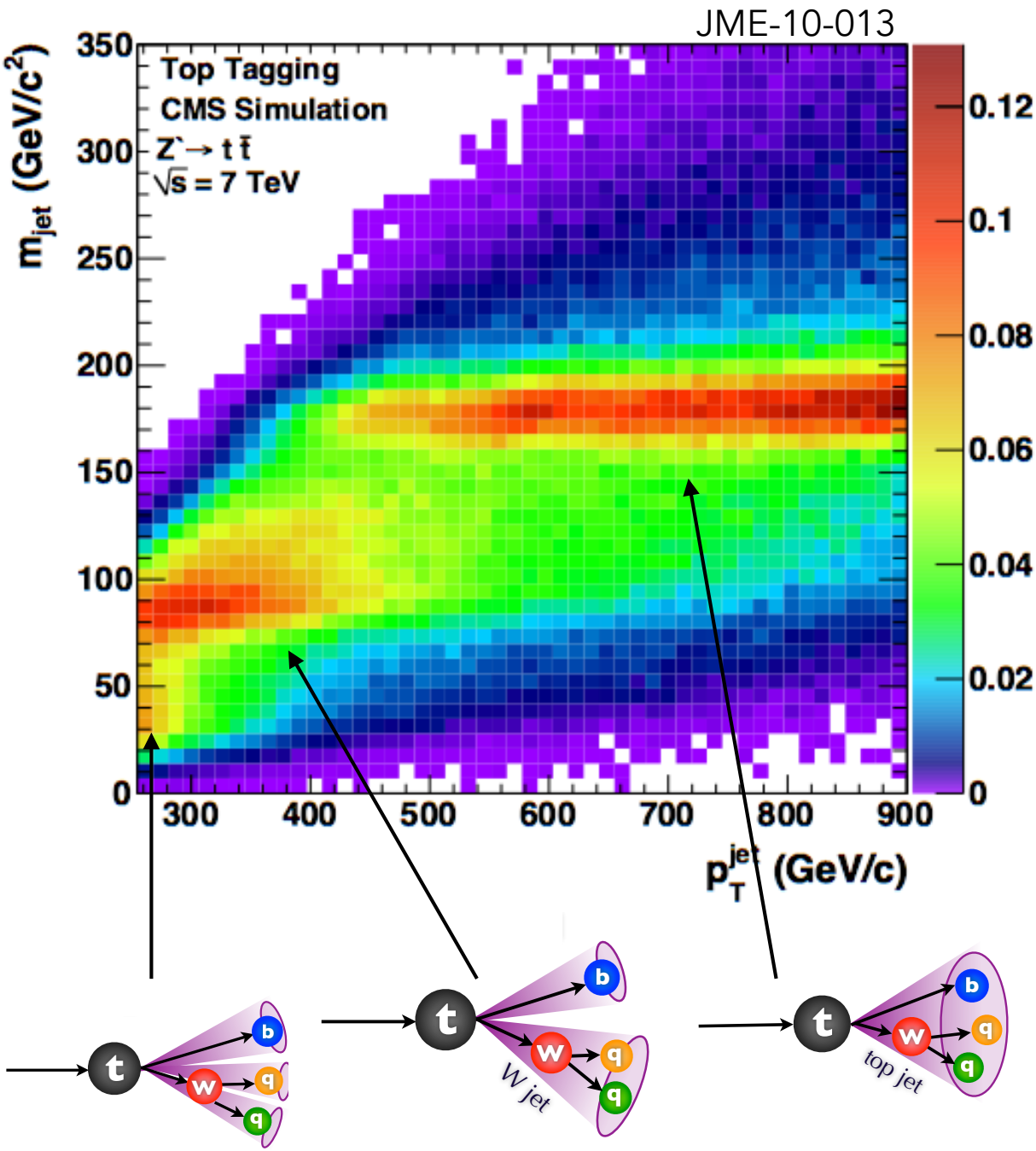


Jet Mass

- For heavy objects, the picture is different
- The jet mass is relatively stable at the heavy particle mass
- Choose correct cone size to reconstruct all decay products in jet:

$$\Delta R \sim \frac{2 \cdot m_{\text{particle}}}{p_T}$$

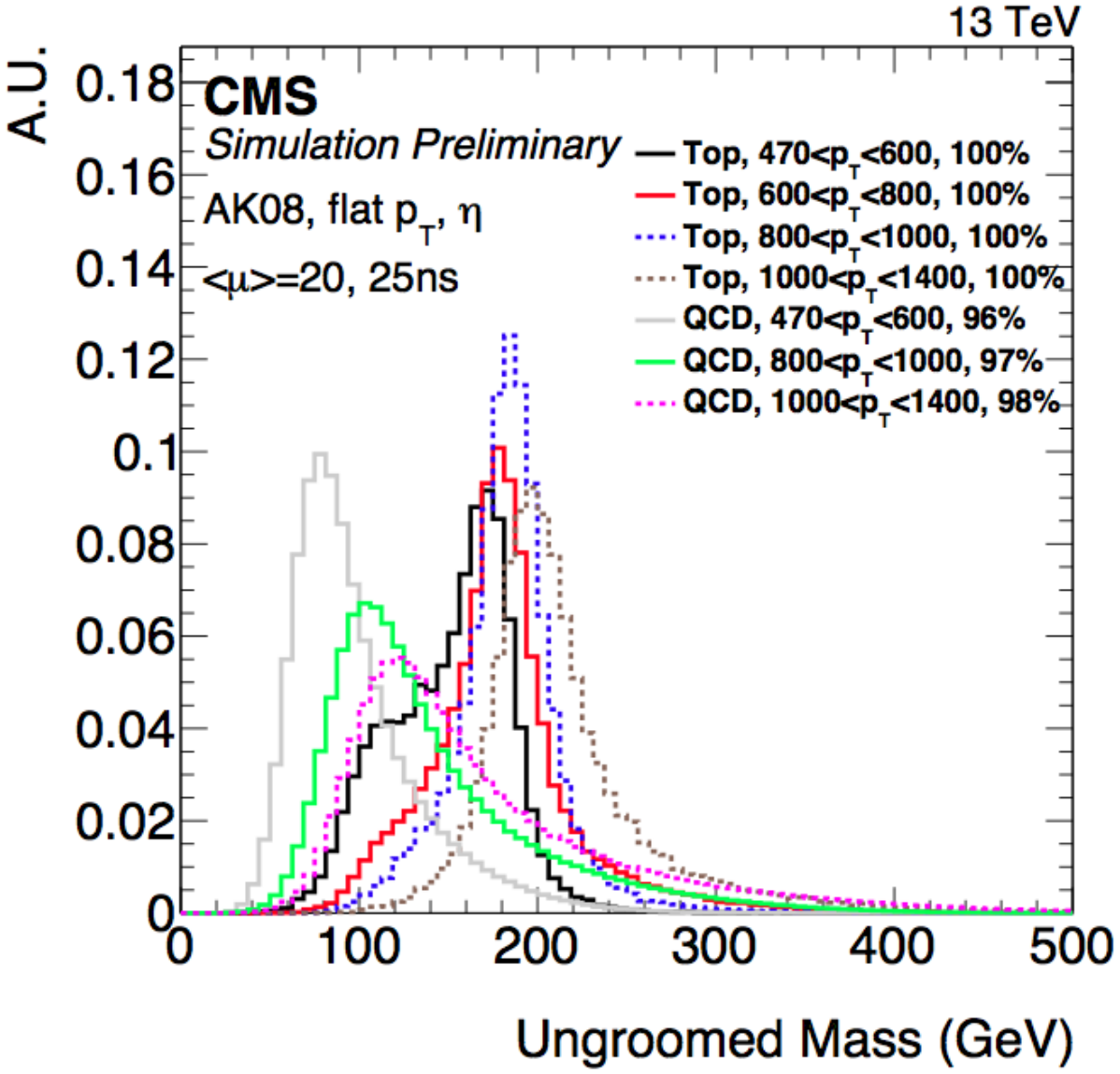
- CMS uses R = 0.8 for heavy object reconstruction
  - Merged W/Z at p_T ~200 GeV
  - Merged top at p_T ~400 GeV



Jet Mass

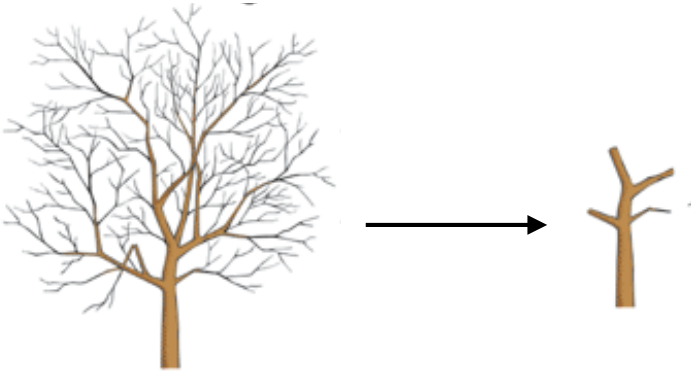
JME-15-002

- Some discrimination obtained when using this ‘raw’ jet mass
- We can do better by looking inside the jet at the individual constituents
- Using **jet grooming** algorithms can improve the discrimination between QCD and top quark jets

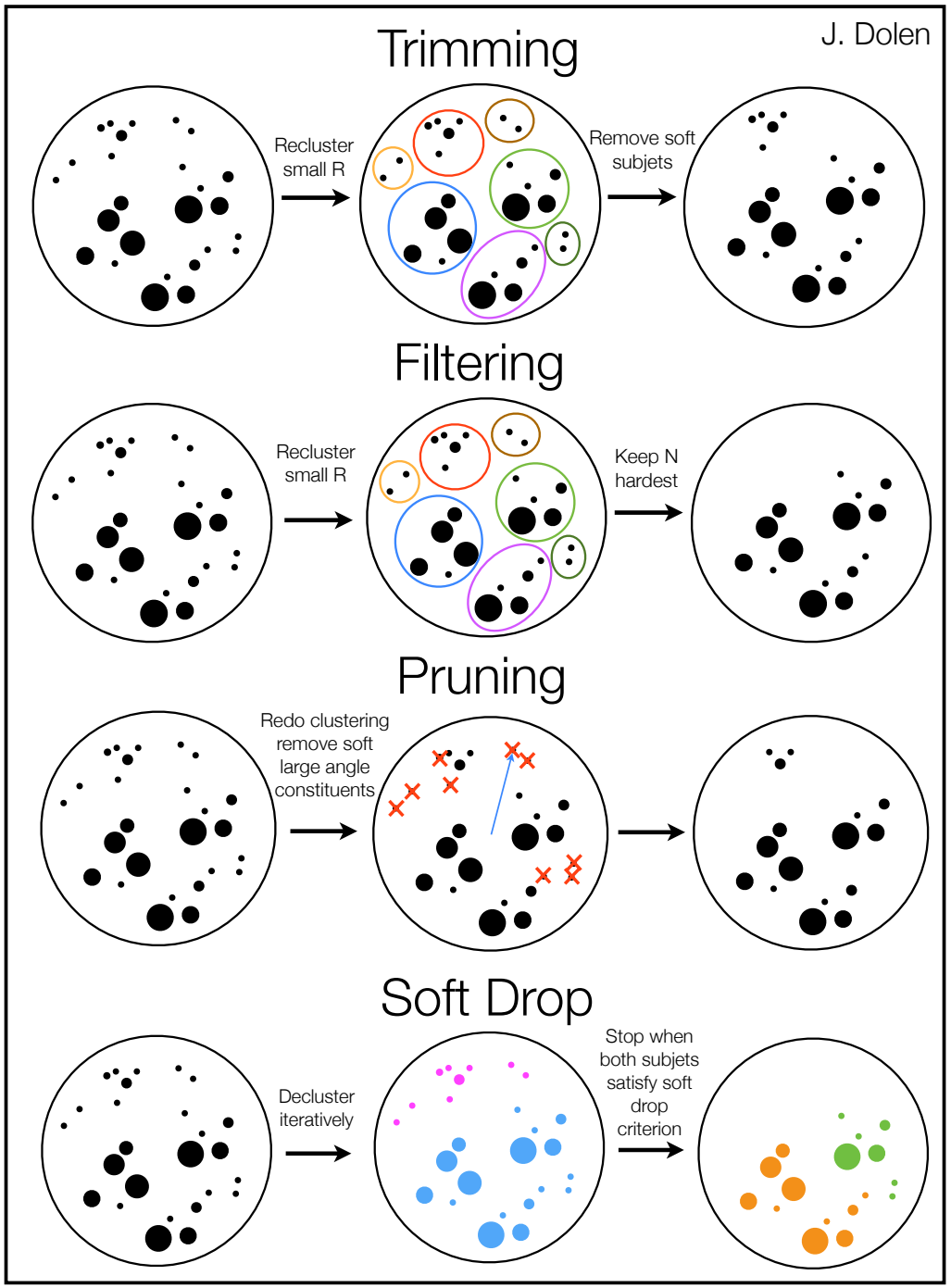


Jet Grooming Algorithms

- Many different approaches
- Standard idea: remove **soft and wide-angle radiation** from within the jet
  - Recluster with smaller R
  - Remove subjets
  - Remove constituents during clustering



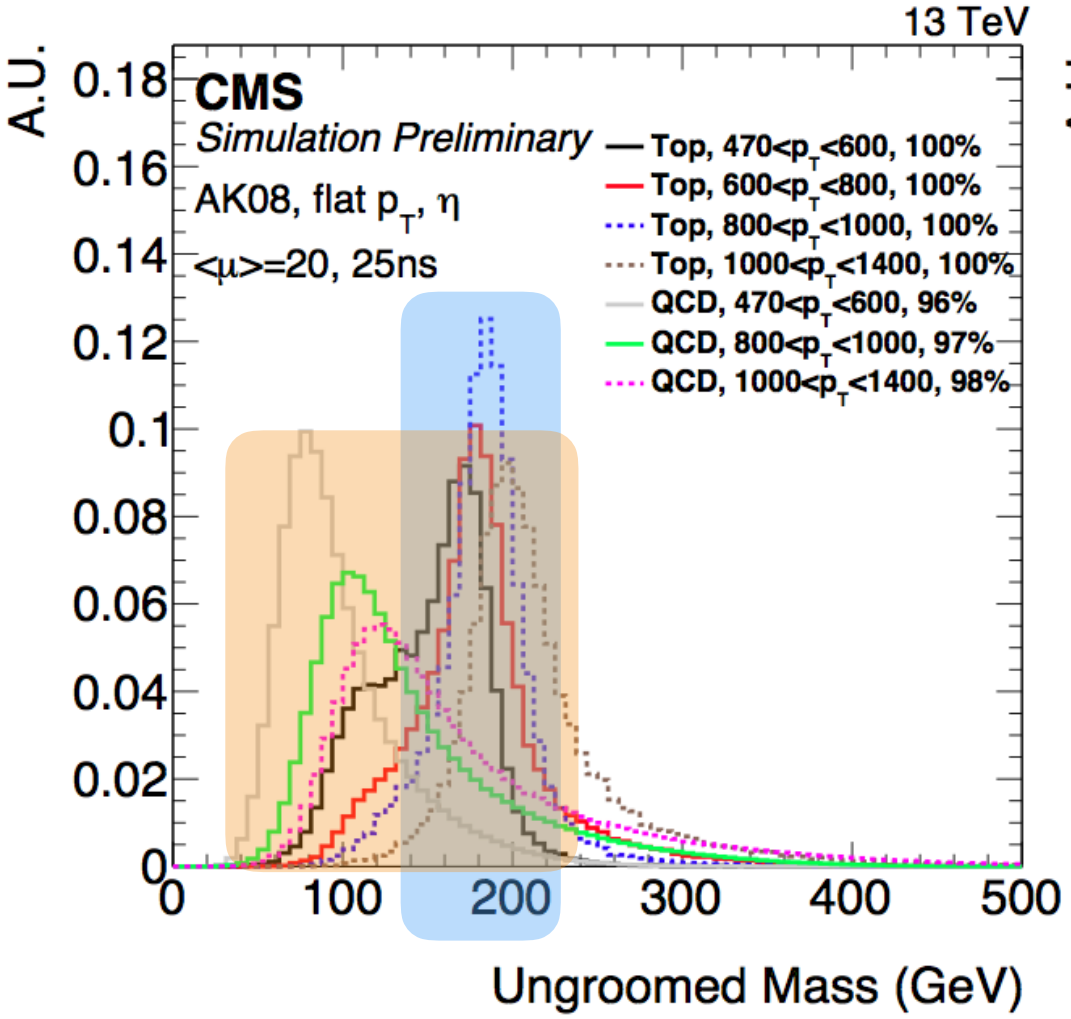
- Soft drop algorithm chosen for latest analyses

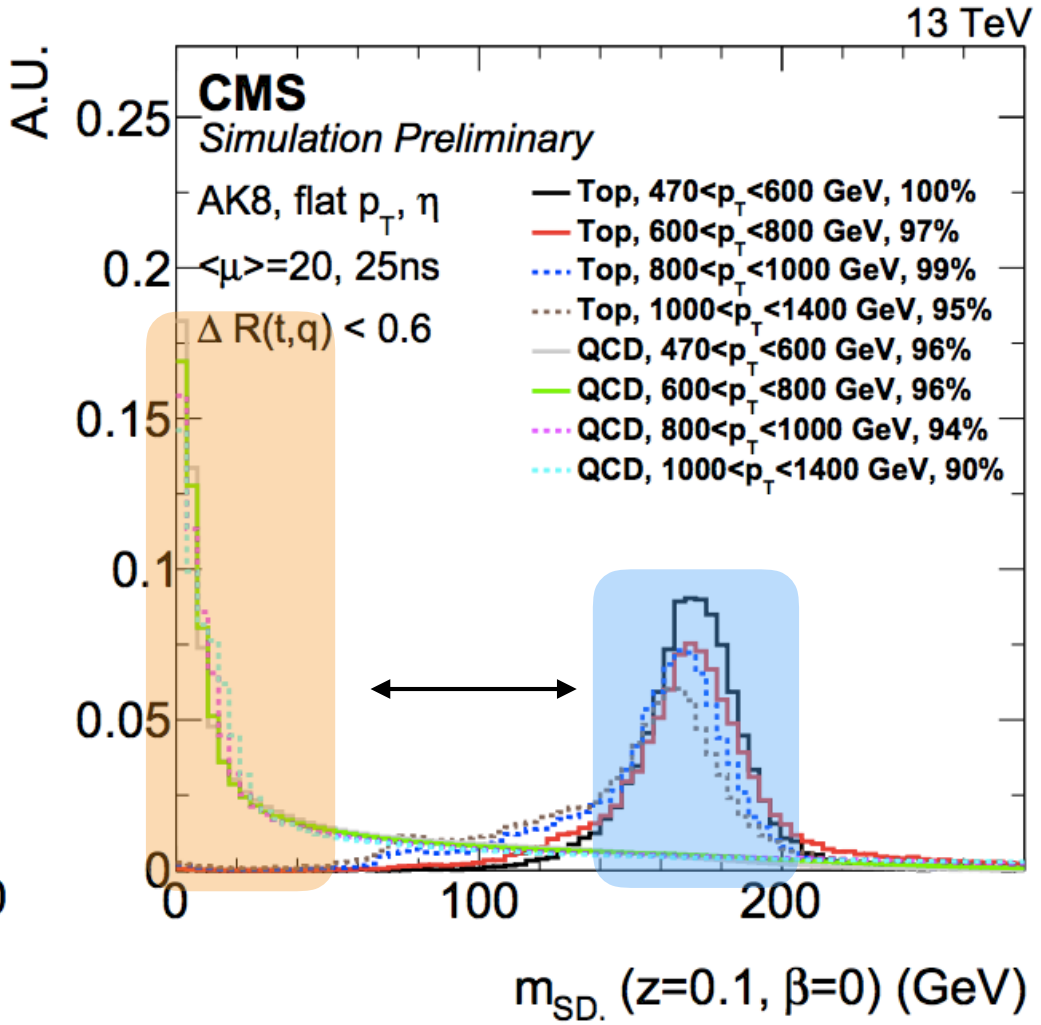


Jet Grooming Algorithms

JME-15-002

- Dramatically improves the separation of QCD and top quark jets
- Merged top quarks can be identified with a window around the top quark mass



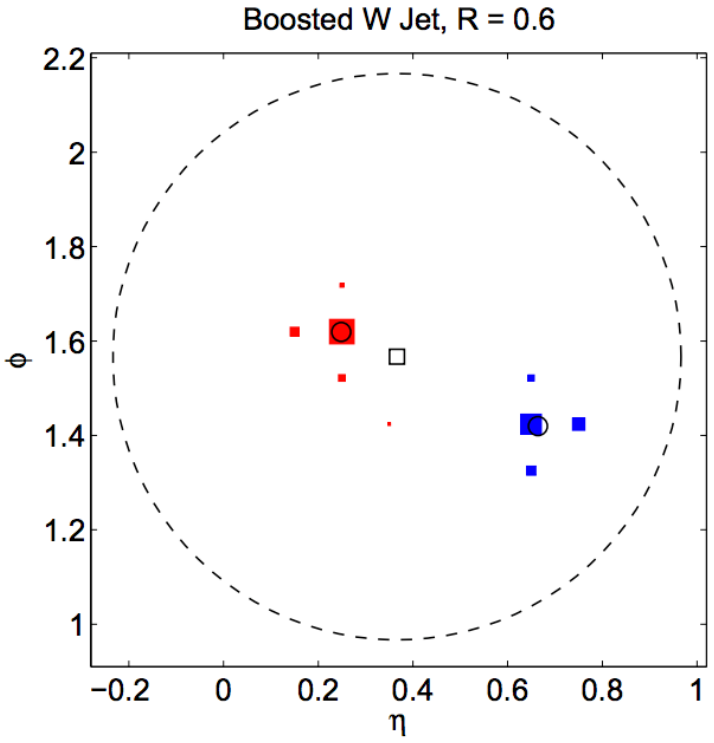


Topological Algorithms

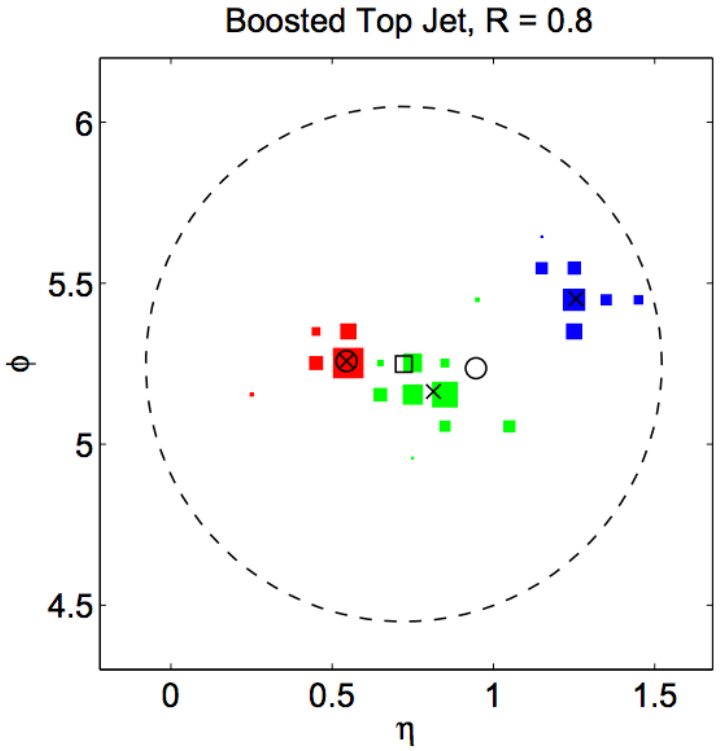
arXiv:1011.2268

- We know how many final state objects to expect from these decays
  - Can look inside the jet for the expected substructure
    - Top decays → 3 subjets
    - W/Z/H decays → 2 subjets
- A quantity called N-subjettiness is a measure of how consistent a jet is with a hypothesized number of subjets
  - Low τ_N → consistent with N (or fewer) subjets
  - Ratios used for additional discrimination

$$\tau_N = \frac{1}{\sum_i p_{T,i} \cdot R} \sum_i p_{T,i} \cdot \min(\Delta R_{1,i}, \Delta R_{2,i}, \ldots, \Delta R_{N,i})$$



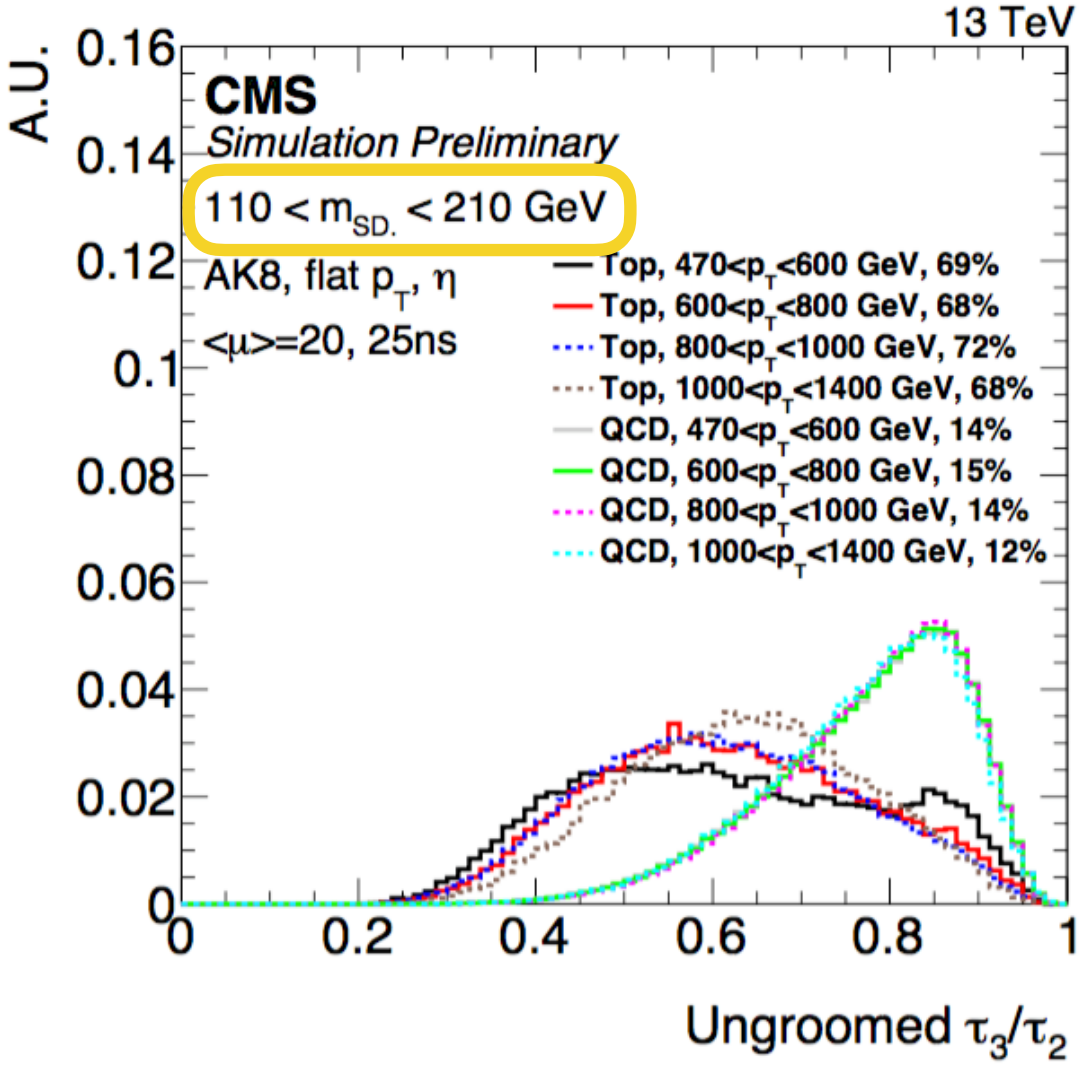
□ 1 subjet hyp.
○ 2 subjet hyp.
× 3 subjet hyp.



N-Subjettiness

JME-15-002

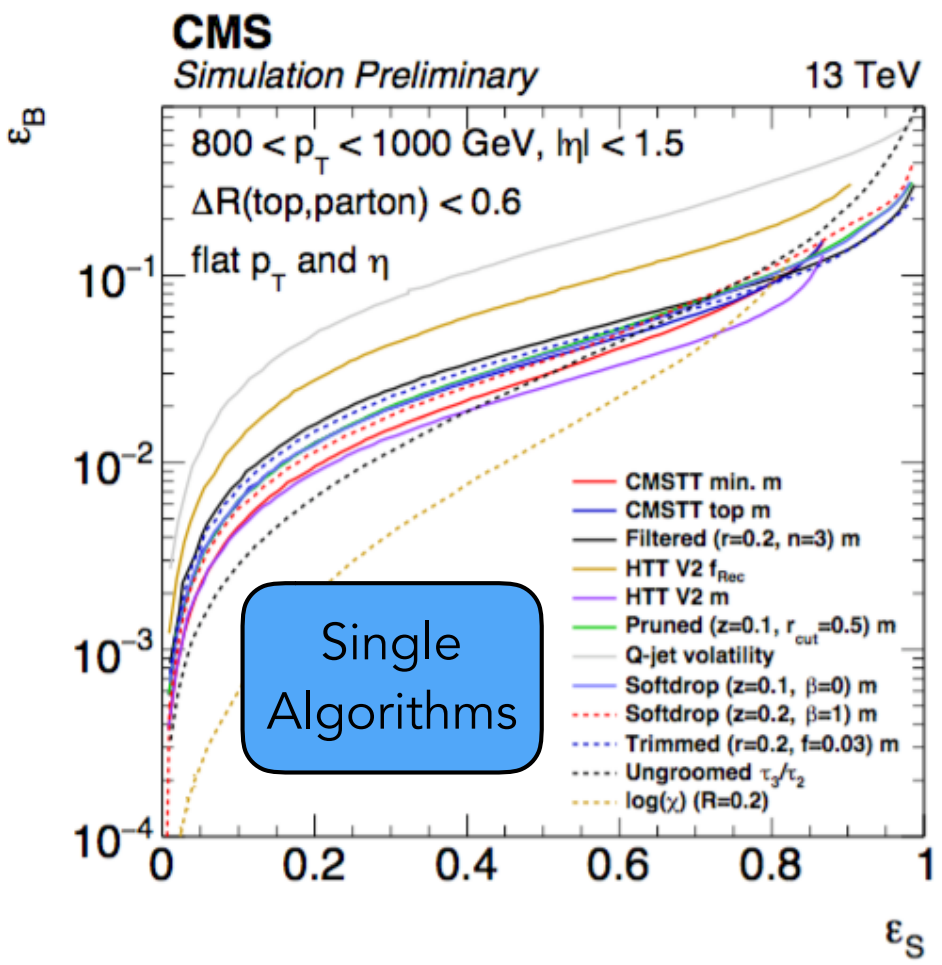
- Ratios used for discrimination
  - τ_3 / τ_2 for top quark jets
  - τ_2 / τ_1 for W/Z/H jets
- Provides additional power when used in conjunction with the groomed jet mass
- **Can we do better?**

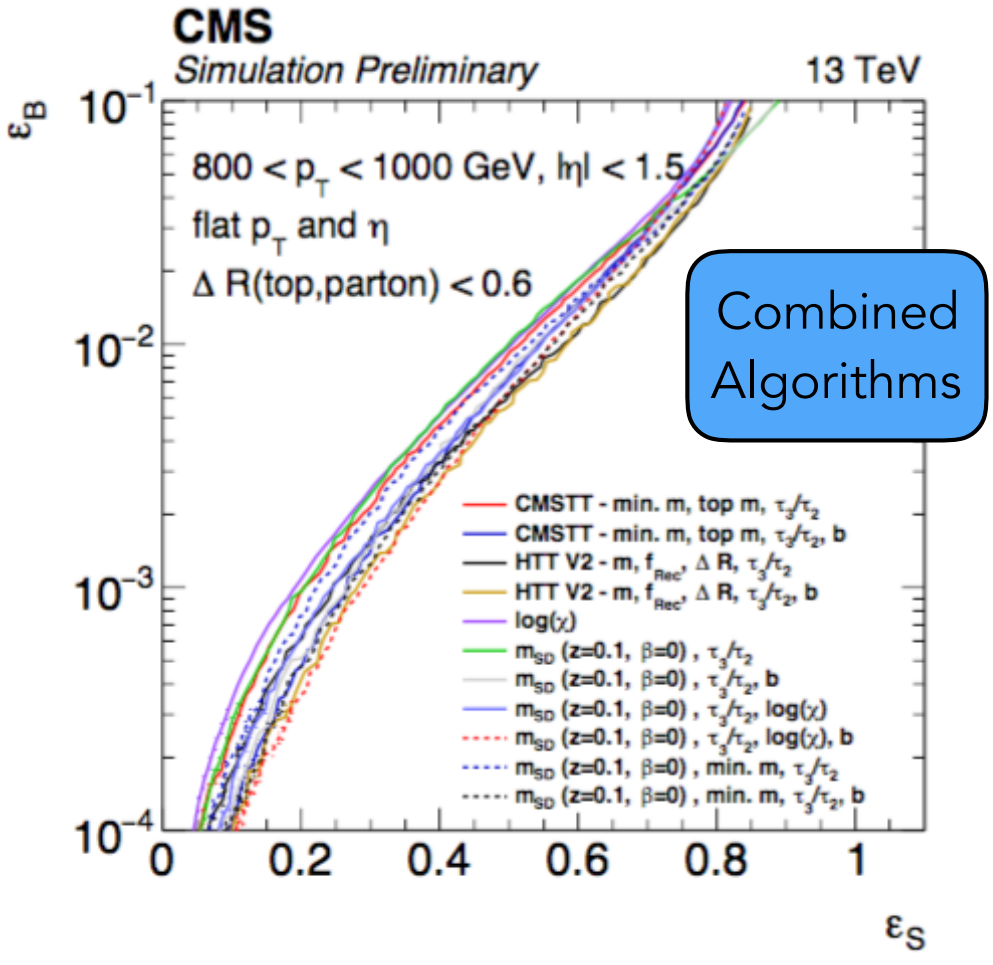


Combining Algorithms

JME-15-002

- The algorithms provide mutual information that increases performance
- Choice of combination — used for top quark identification in 13 TeV analyses:
  - **Soft-drop mass + N-subjettiness (+ b-tagging)**



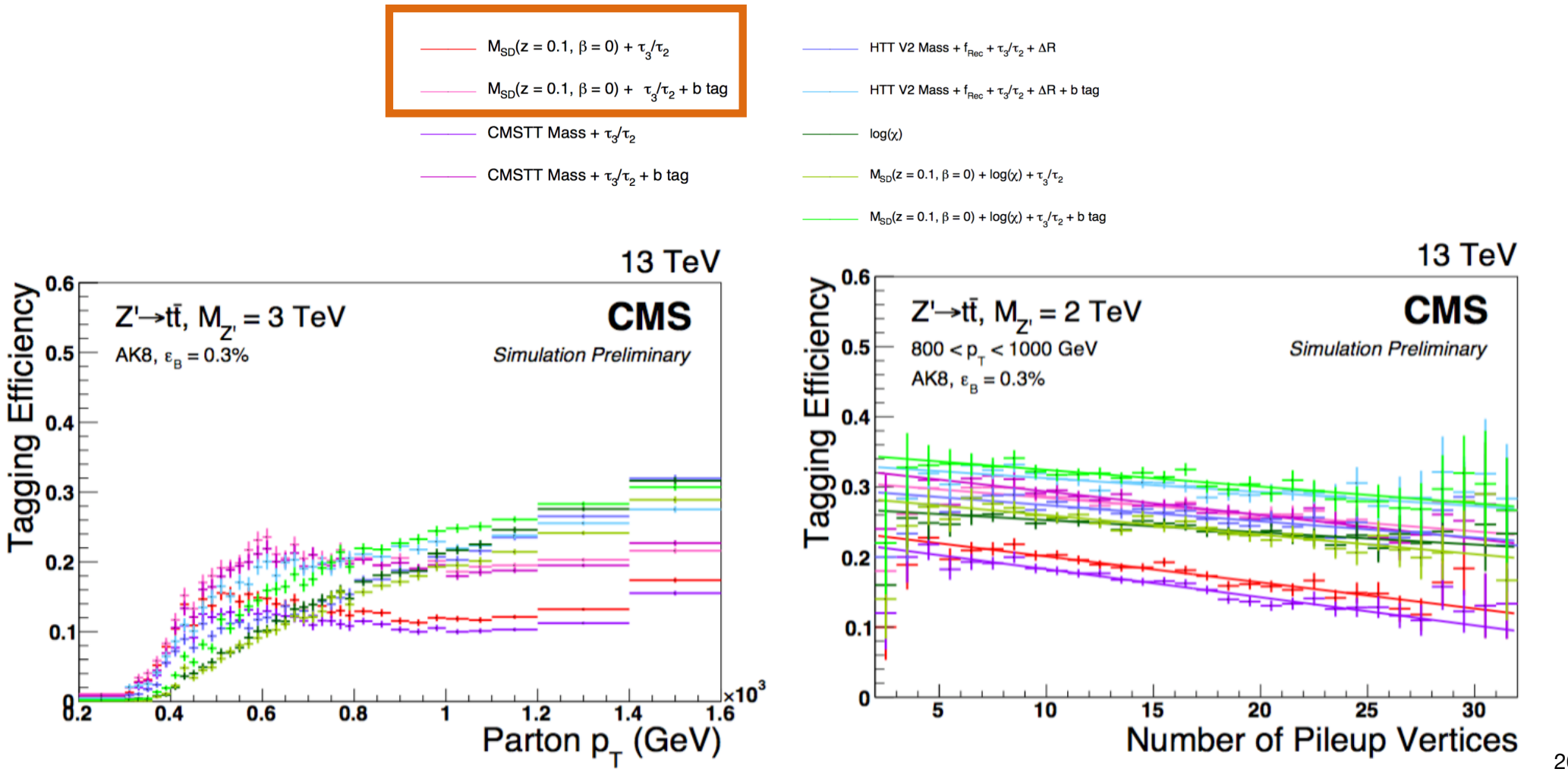


Working Points

JME-15-002

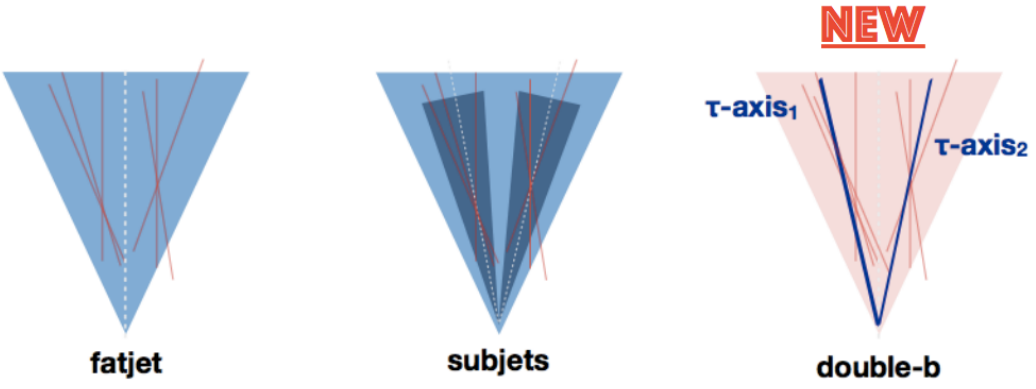
- Choose cut values to ensure stability across both p_T and pileup activity **for a chosen background level** (here 0.3%)
  - m_{SD} in [110, 210] GeV
  - $\tau_{32} < 0.69$

  - Slight degradation in efficiency as a function of pileup



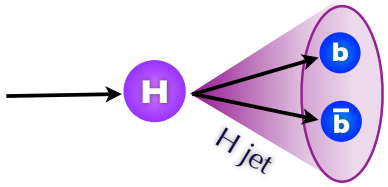
Beyond Hadronic Top Decays

- Similar developments for other boosted heavy object identification

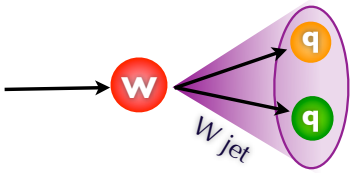


C. Vernieri

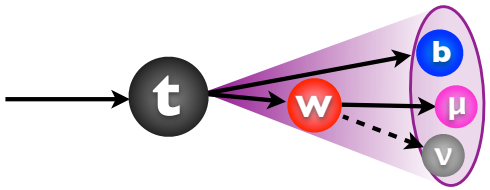
- **Higgs tagging**
  - Look for massive jets with 2 b-tagged subjets

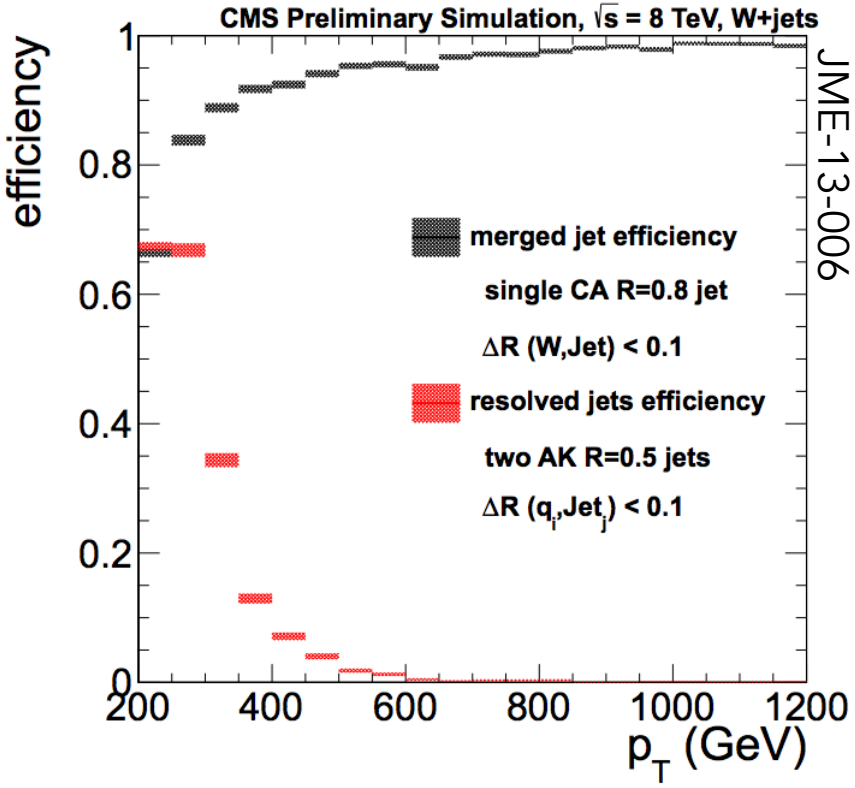


- **W/Z tagging**
  - Use τ_2 / τ_1 for 2-prong decay mode
  - Single-jet reconstruction more efficient at high p_T



- **Leptonic top quark decays**
  - Non-isolated leptons

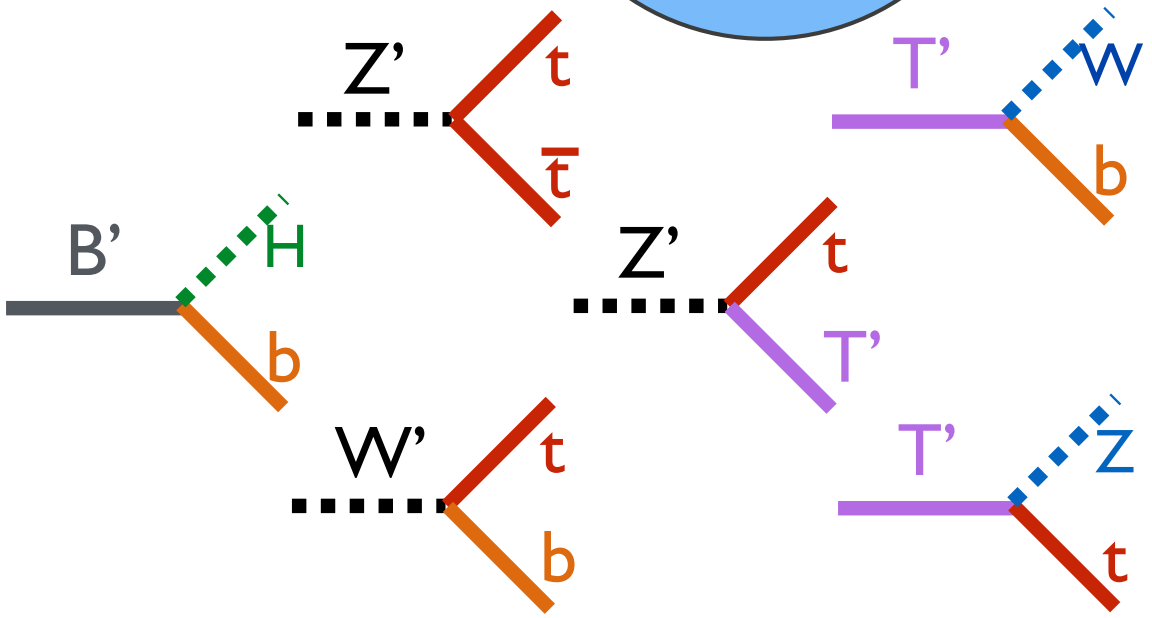




Searches with Jet Substructure

CMS B2G Group

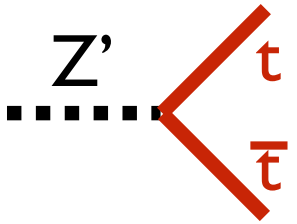
- The 'Beyond 2 Generations' Physics Analysis Group in CMS is focused on searches which decay to massive SM objects (t/W/Z/H)
- **Our specialty is jet substructure and heavy object tagging!**
- There are many searches which take advantage of these signatures
  - Top pair resonances
  - Top + bottom resonances
  - Diboson resonances
  - Heavy vector-like quarks
  - Hybrid decay modes

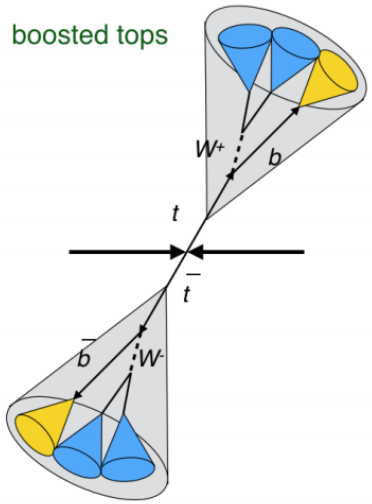


Top Quark Pair Resonances

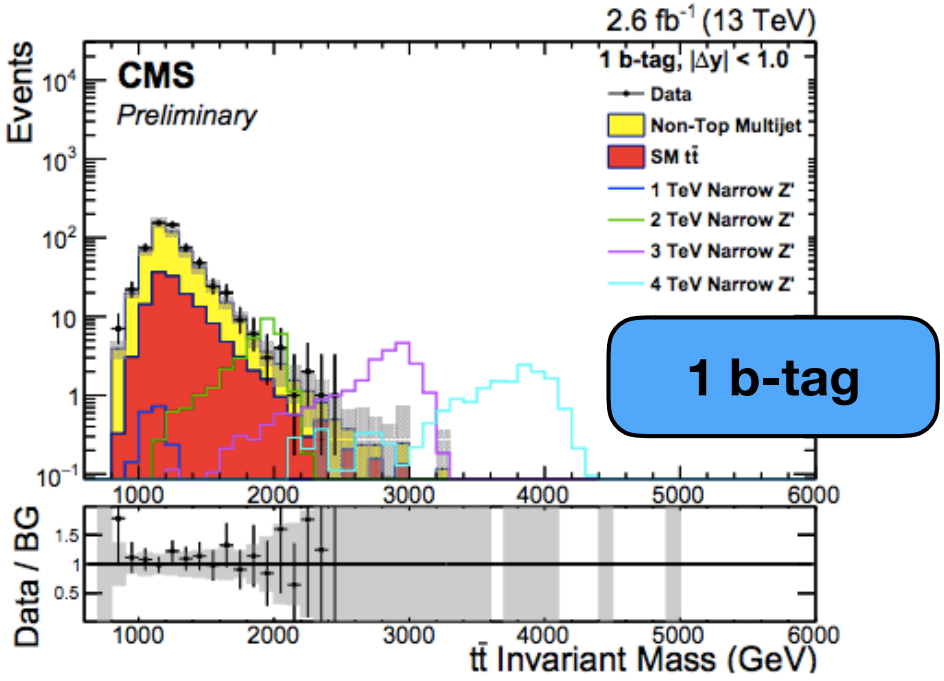
B2G-15-003

- Hadronic channel — dijet event topology

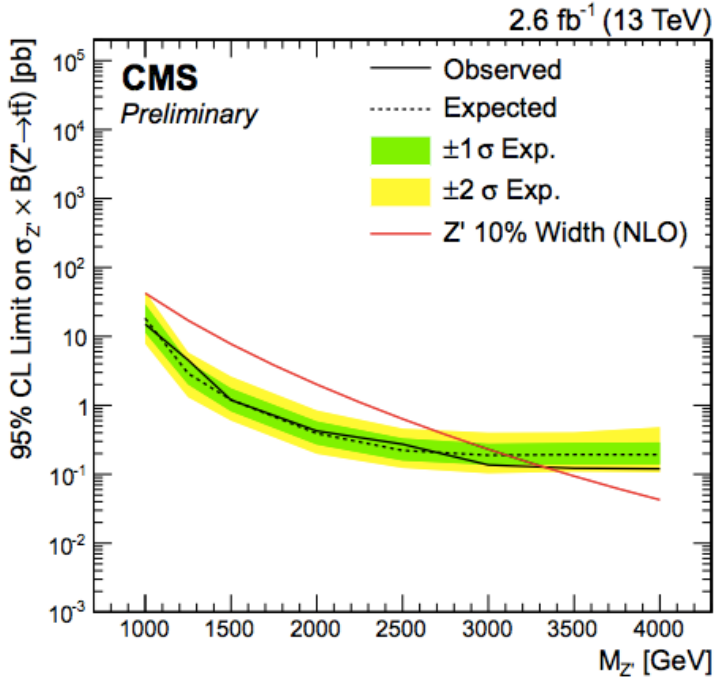




- Search for events with 2 back-to-back top-tagged jets
  - p_T > 400 GeV
  - m_{SD} in [110, 210] GeV
  - τ_{32} < 0.69



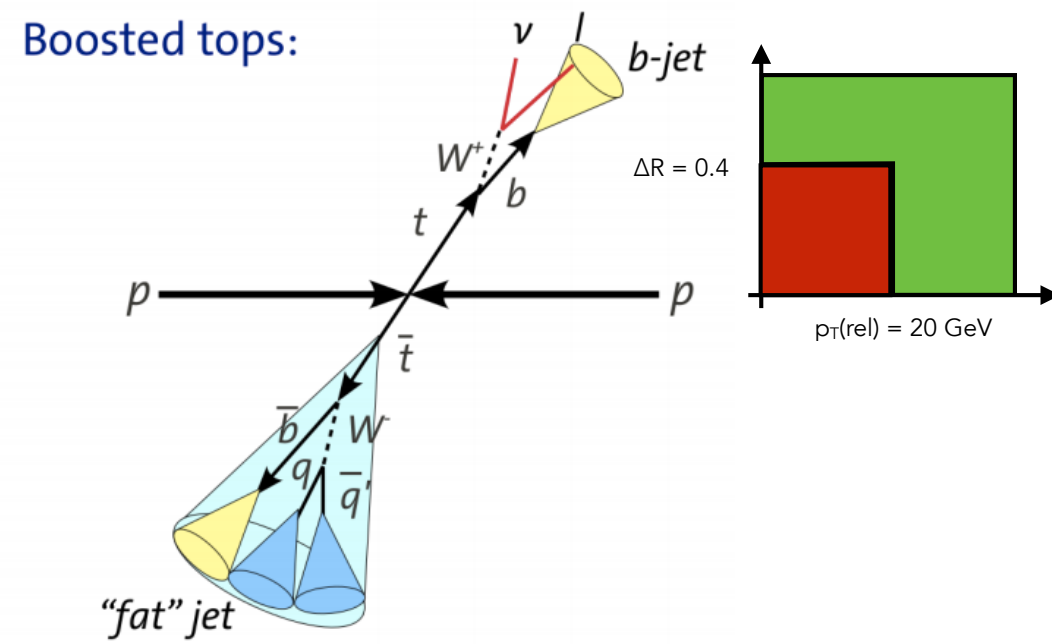
- Divide events into categories based on number of subjet-b-tagged jets and rapidity separation
- Use top pair invariant mass to discriminate signal

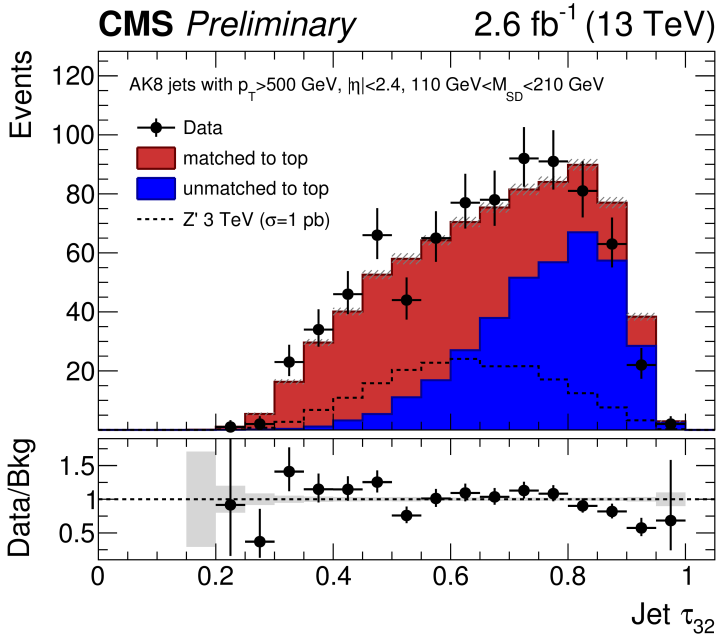


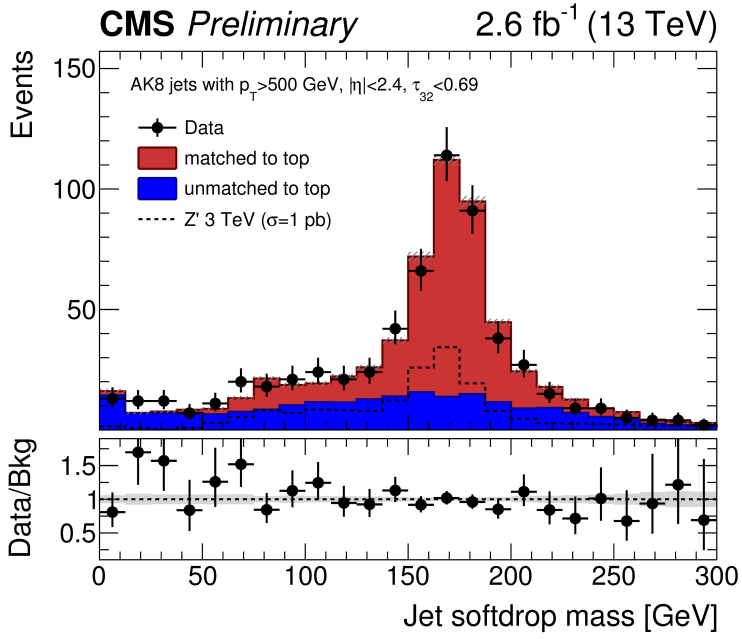
Top Pair Resonances

B2G-15-002

- Also consider the lepton+jets decay mode in this search
  - At high p_T, lepton may overlap with jet → special reconstruction needed to reject QCD
    - $\Delta R > 0.4$ or $p_T(rel) > 20$ GeV
- Same top-tagging algorithm used

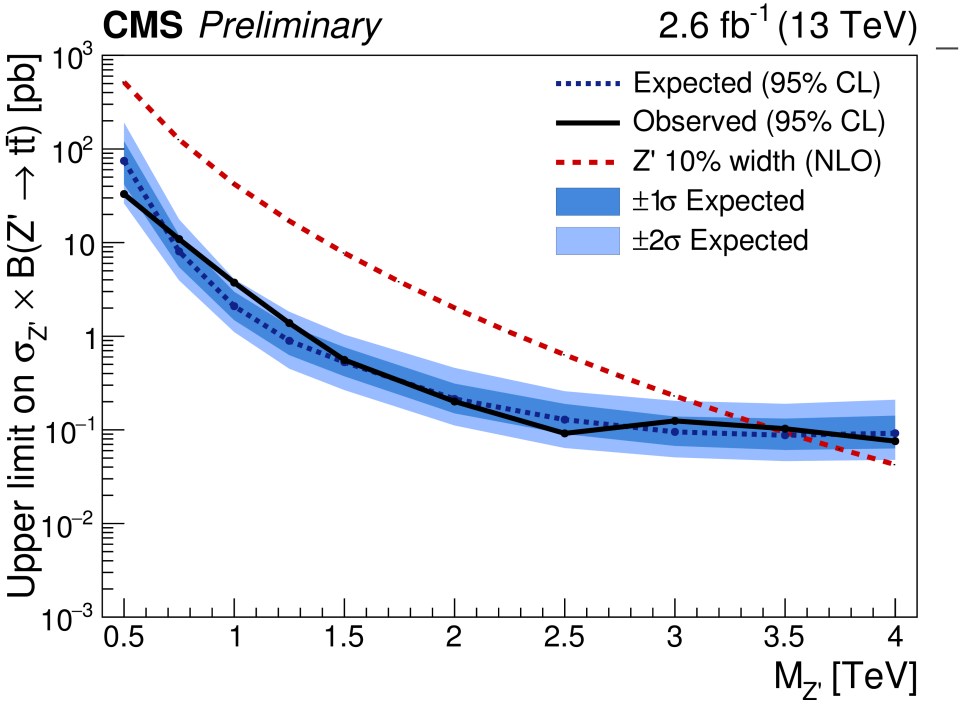


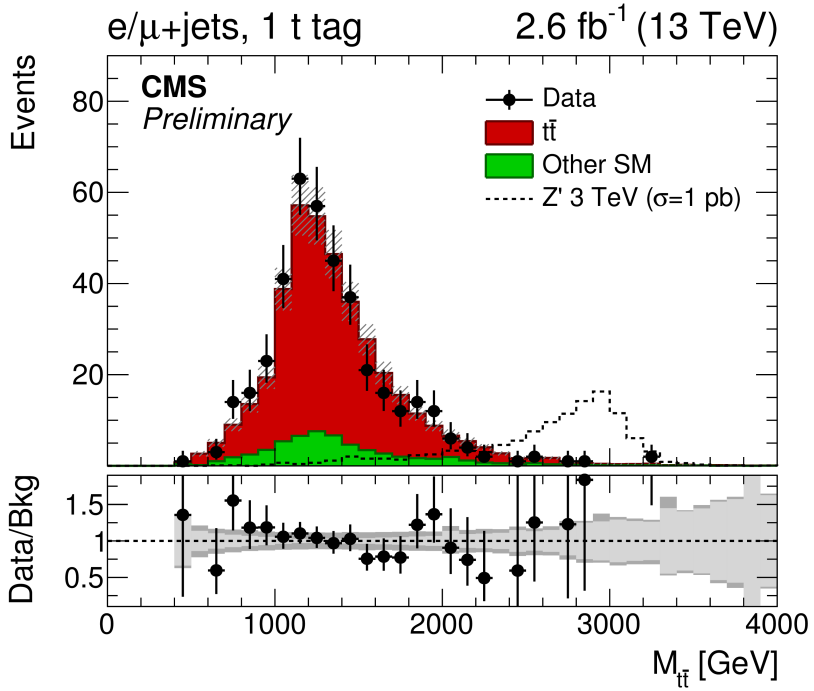


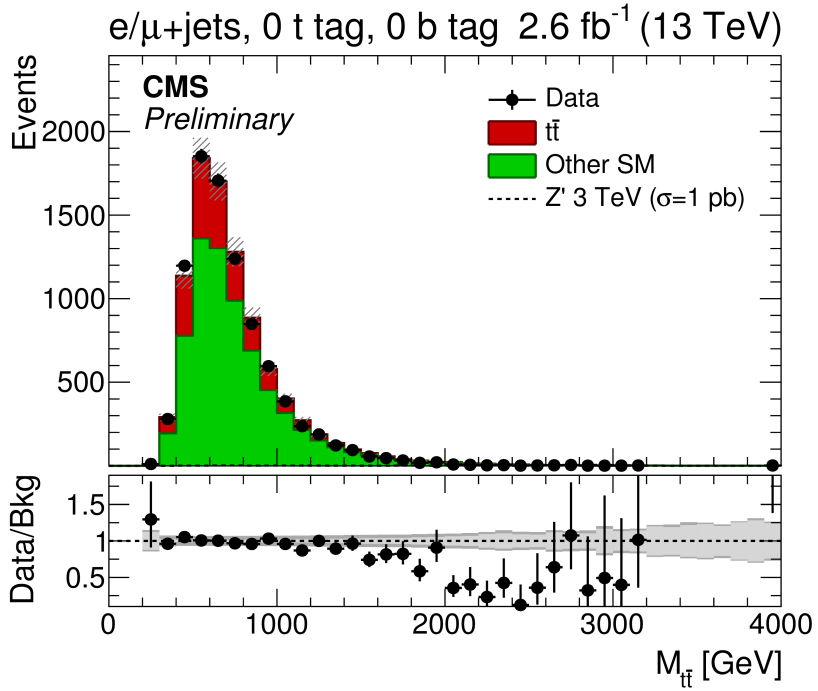


Top Pair Resonances

- Events are divided by lepton flavor and number of b/t-tagged jets
  - Top pair invariant mass used for discrimination
- Limits range from 2.3 - 4.0 TeV depending on the physics model



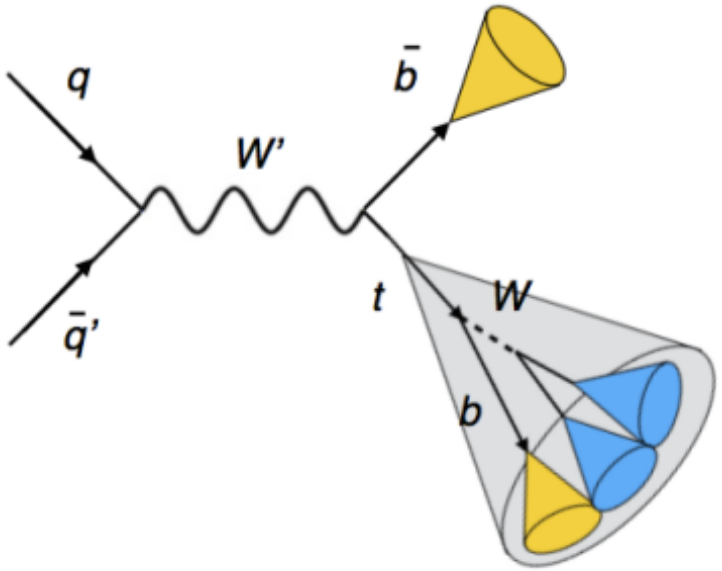


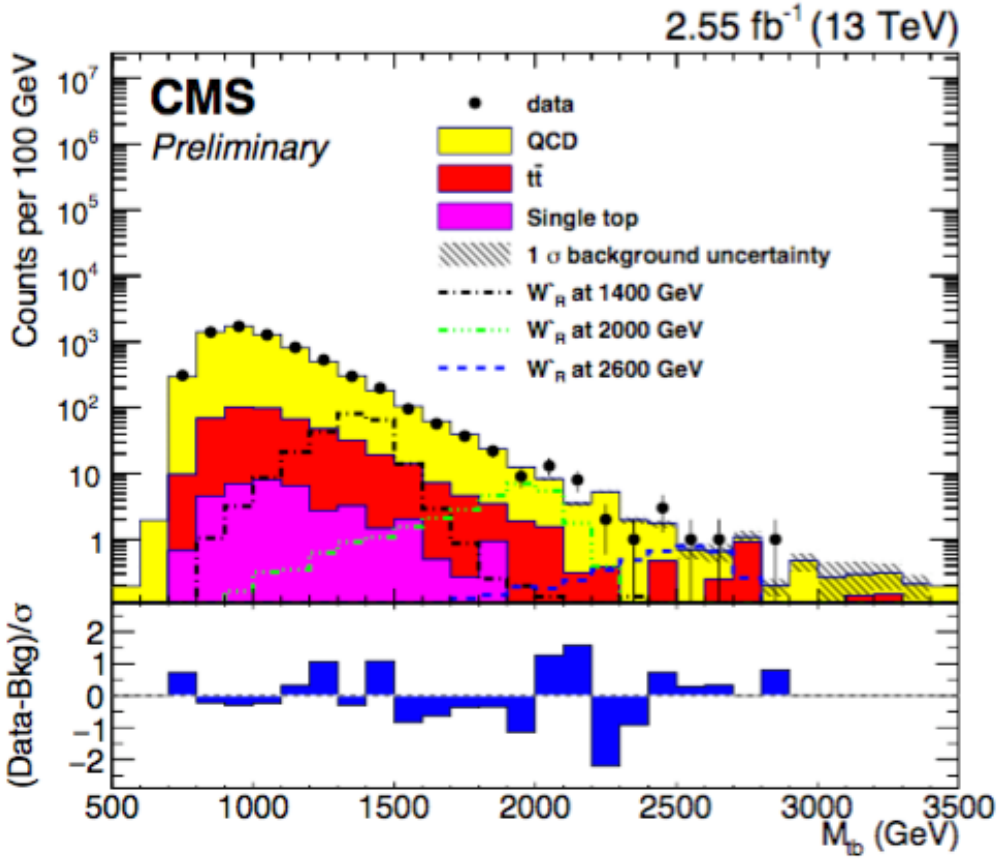


t+b Resonances

B2G-16-009

- The hadronic channel takes advantage of top-tagging algorithms
- Dijet topology
  - Top-tagged jet opposite from a high-pT b-jet
  - Each jet p_T > 350 GeV
  - b-jet mass < 70 GeV to remove ttbar background
- QCD background estimated using data-driven mistag rate
- Dijet (t+b) mass used for signal discrimination



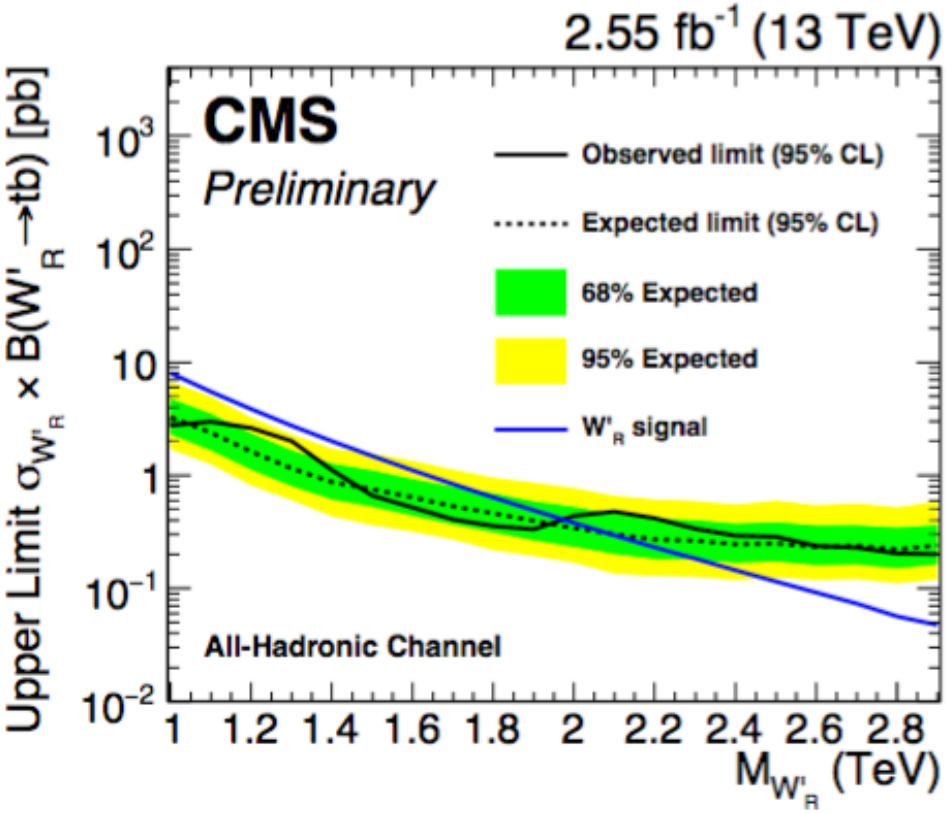


W′ Search Results

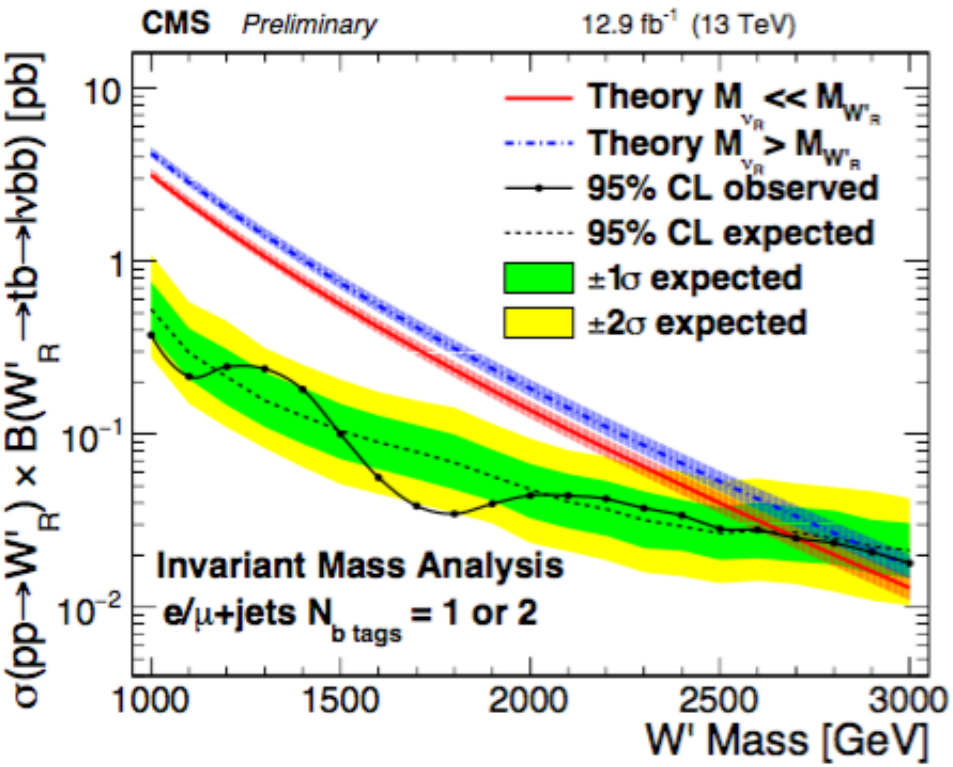
B2G-16-009
B2G-16-016

- Mass exclusions up to 2.0 TeV (hadronic) and 2.7 TeV (lepton+jets)
- Combination currently in progress with updated datasets

Hadronic

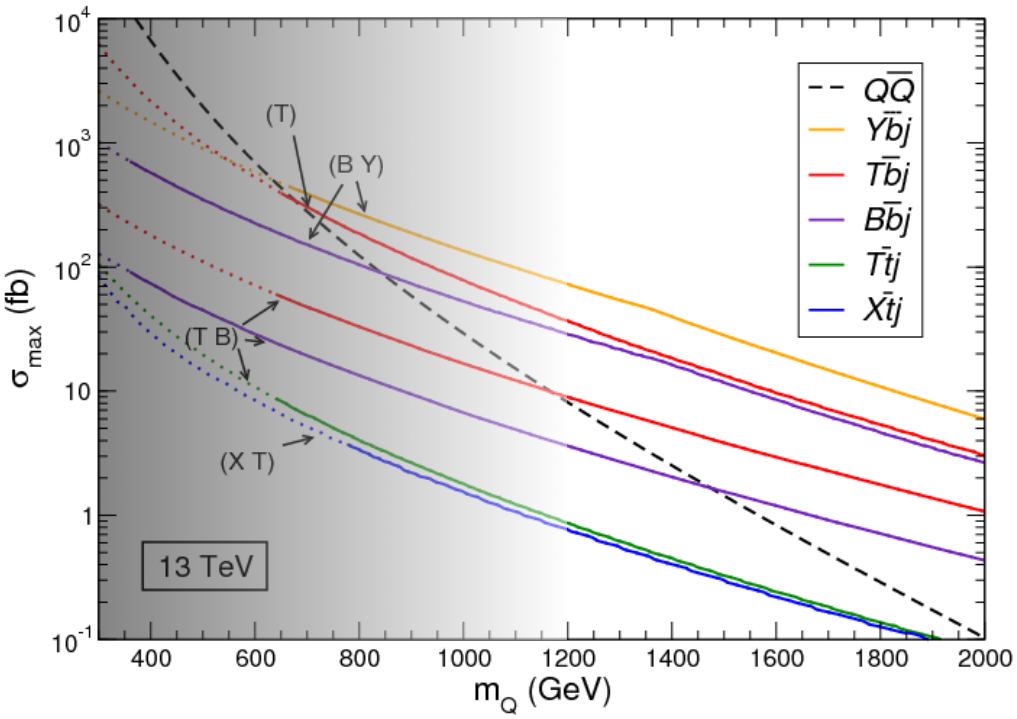


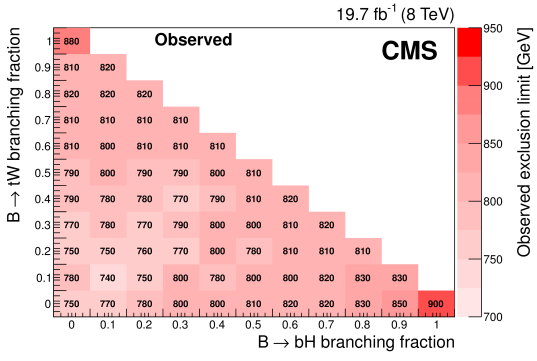
Lepton+Jets

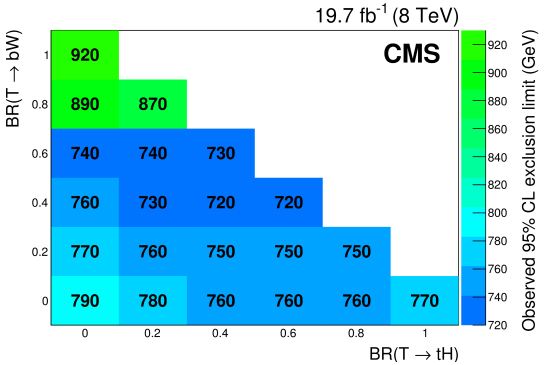


Vector-Like Quarks

- A special type of 4th generation particle predicted by many models
  - Non-typical Higgs coupling
  - Solves hierarchy problem while escaping Higgs cross section constraints
- Different decay modes possible
  - B → tW, bZ, bH
  - T → bW, tZ, tH
  - **With high p_T T/B — heavy reliance on jet substructure techniques for reconstruction!**
  - *"Examine triangles, not points"*
- Diverse final states possible when considering pair production
  - Single production also important at 13 TeV



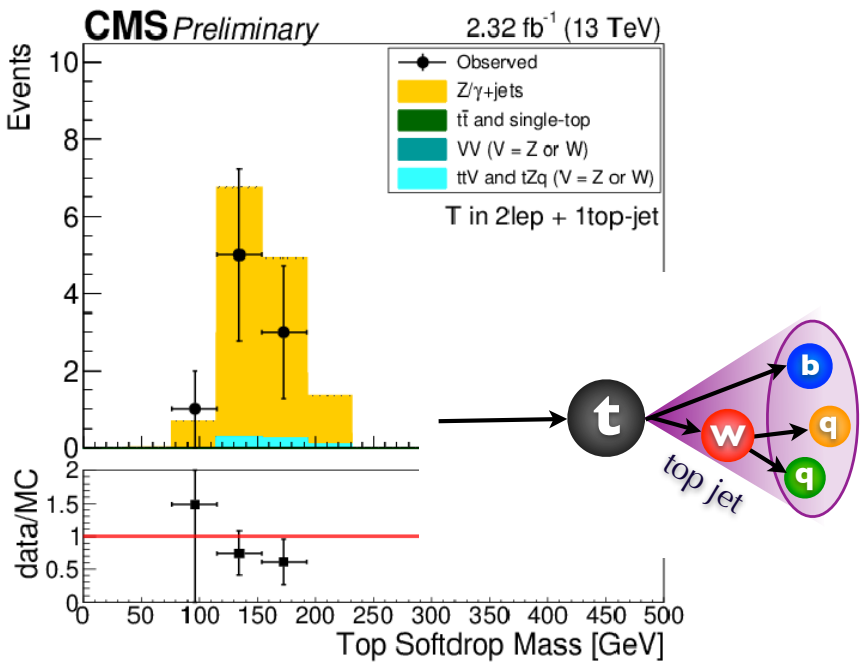


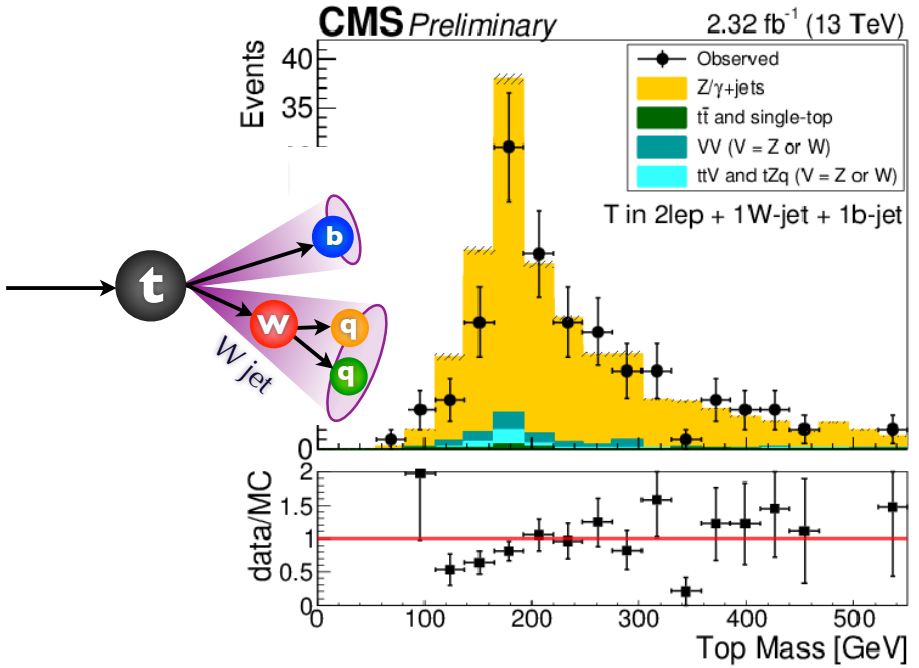


B2G-13-005
B2G-13-006

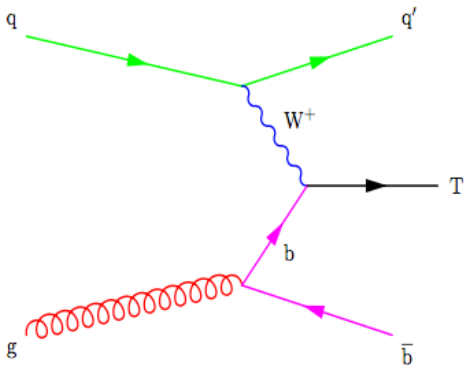
New VLQ Search Results

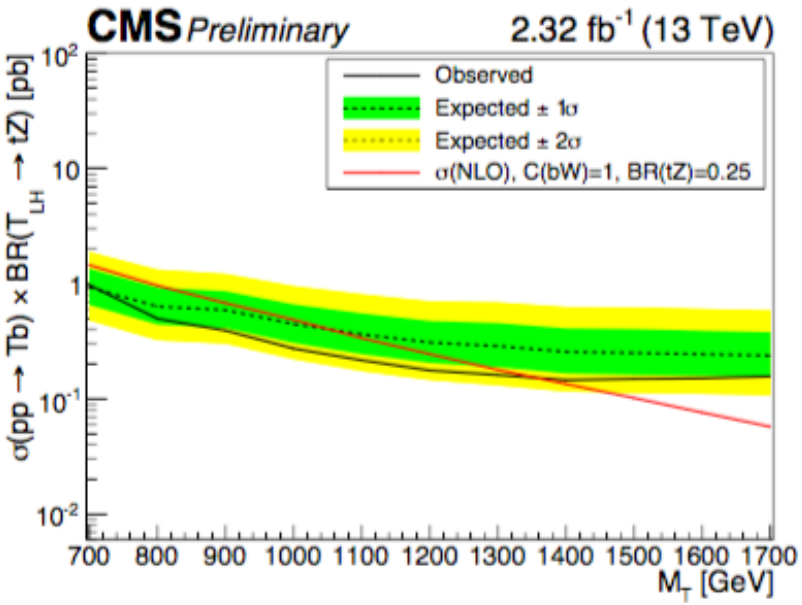
B2G-16-001





- One recent analysis: T → tZ in the dilepton channel
  - Top mass reconstruction using substructure
    - W-tag + b-jet
    - Fully merged top jet

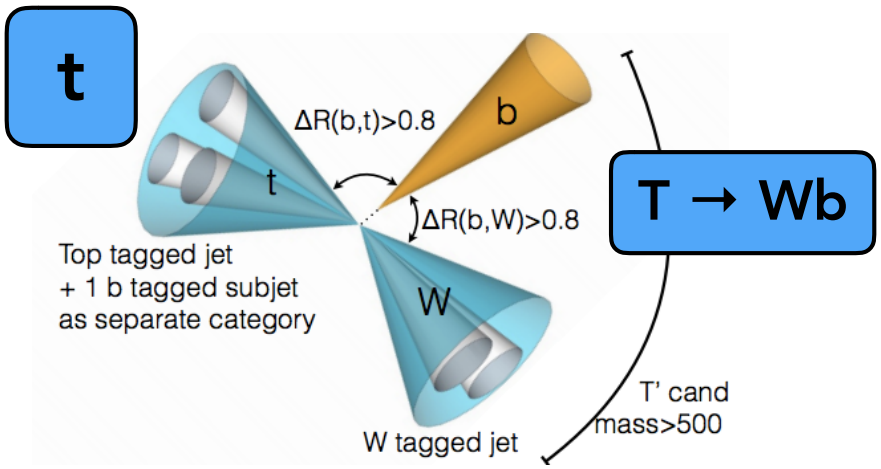




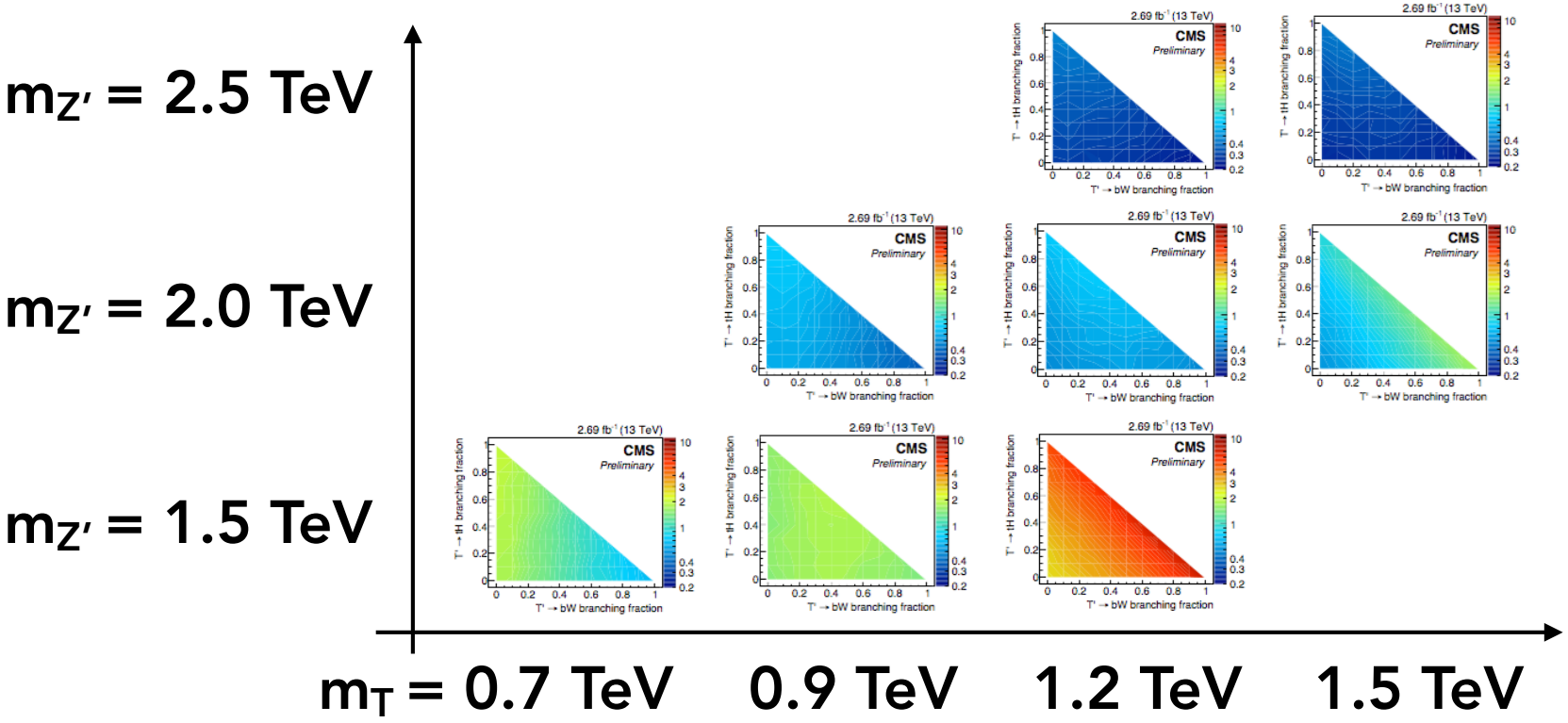
Hybrid Analyses

B2G-16-013

- There are models which allow the new heavy gauge bosons to decay to the VLQs
  - Z′ → tT, bB
  - W′ → bT, tB
  - Masses of Z′, T/B can vary
    - *"Triangles within triangles"*

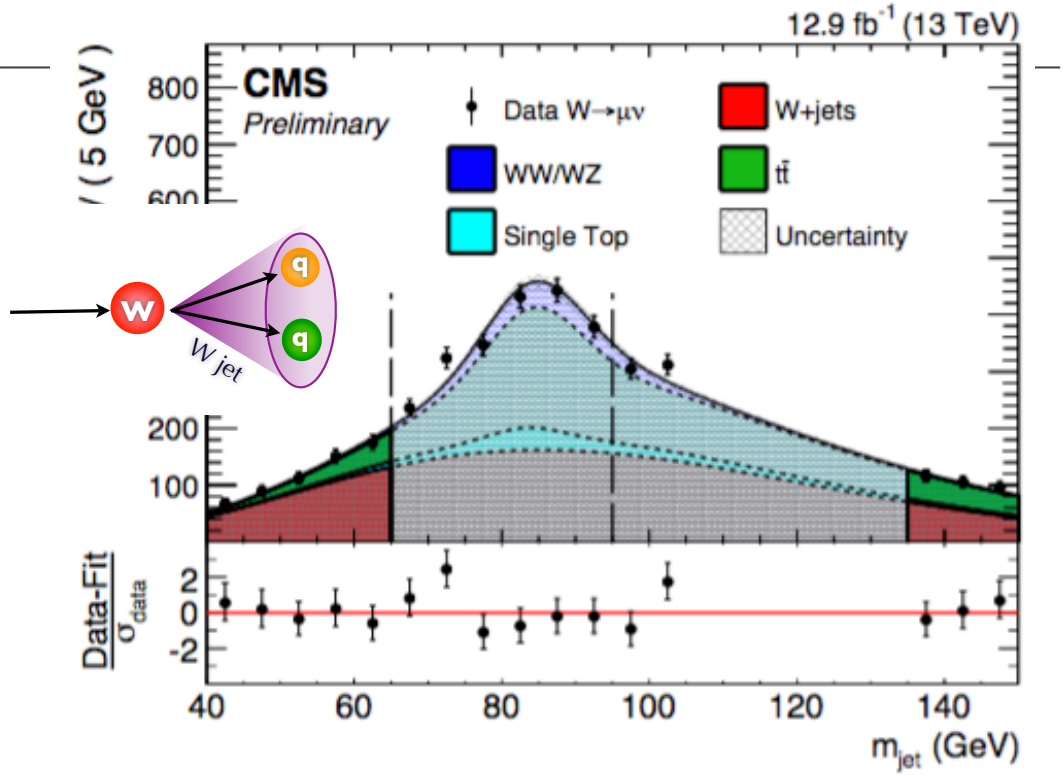


- First public result is search for Z′→ tT



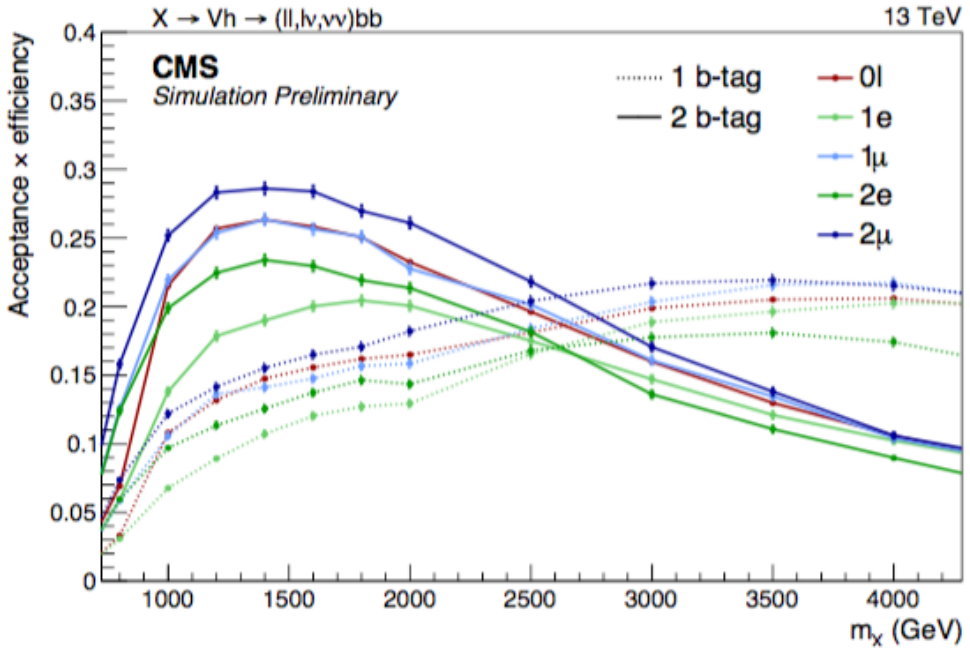
Diboson Resonances

- B2G group also includes searches for heavy resonances decaying to dibosons (WW/WZ/ZZ/HH)
  - **Many opportunities to deploy boosted object tagging with jet substructure**



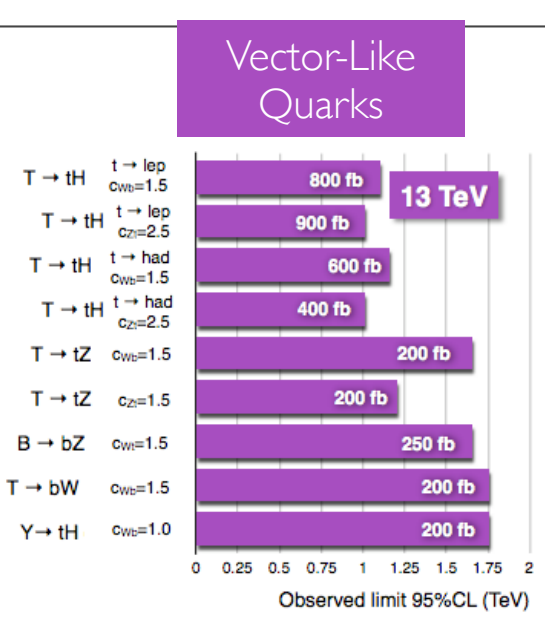
- Some channels use groomed jet mass to identify merged W/Z jets

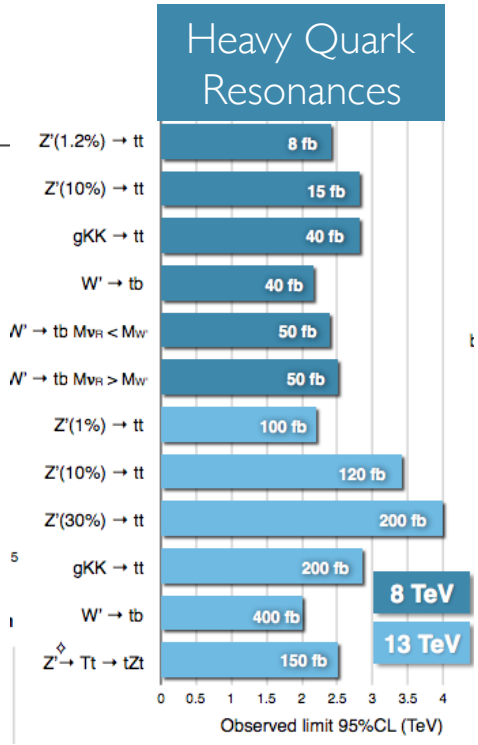
- Boosted Higgs(→bb) bosons better identified at high-mass with a single b-tag due to subjet merging!

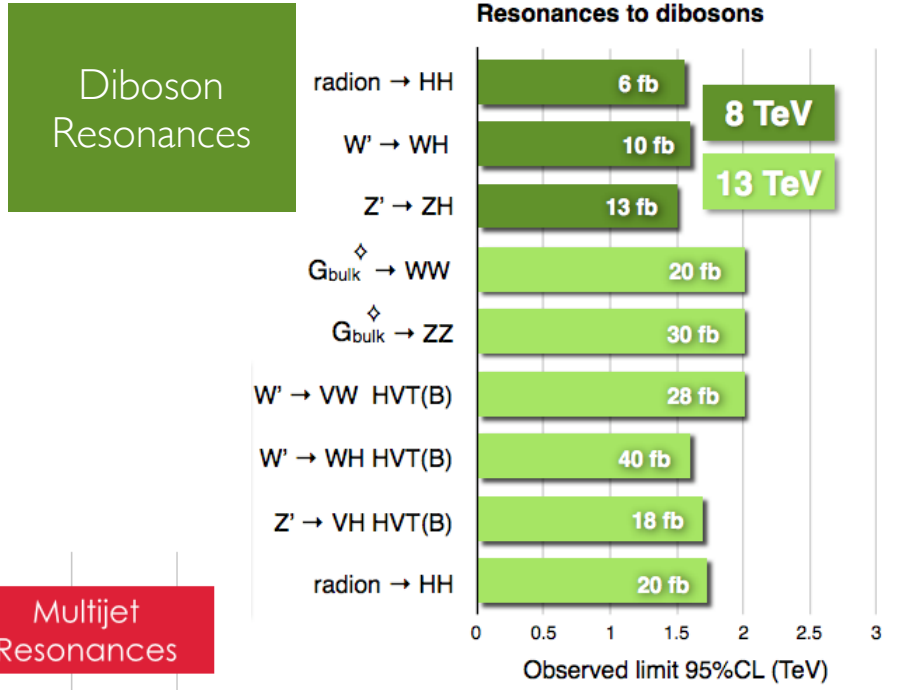


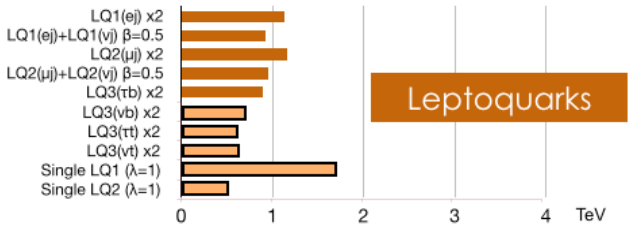
Searches Summary

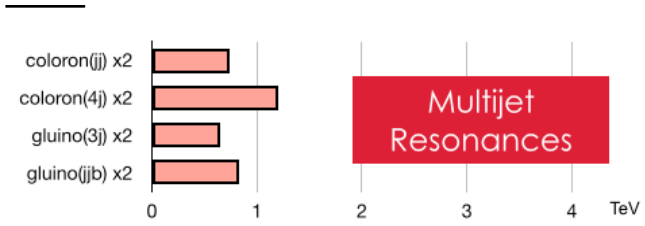
- Many recent search results pushing the several-TeV limits
- Probing higher mass scales means decay products have higher p_T
  - Need jet substructure techniques now more than ever
- Running conditions at the LHC continue to become more and more challenging
  - High pileup
- Is there any hope??
  - YES — we are prepared for discovery!







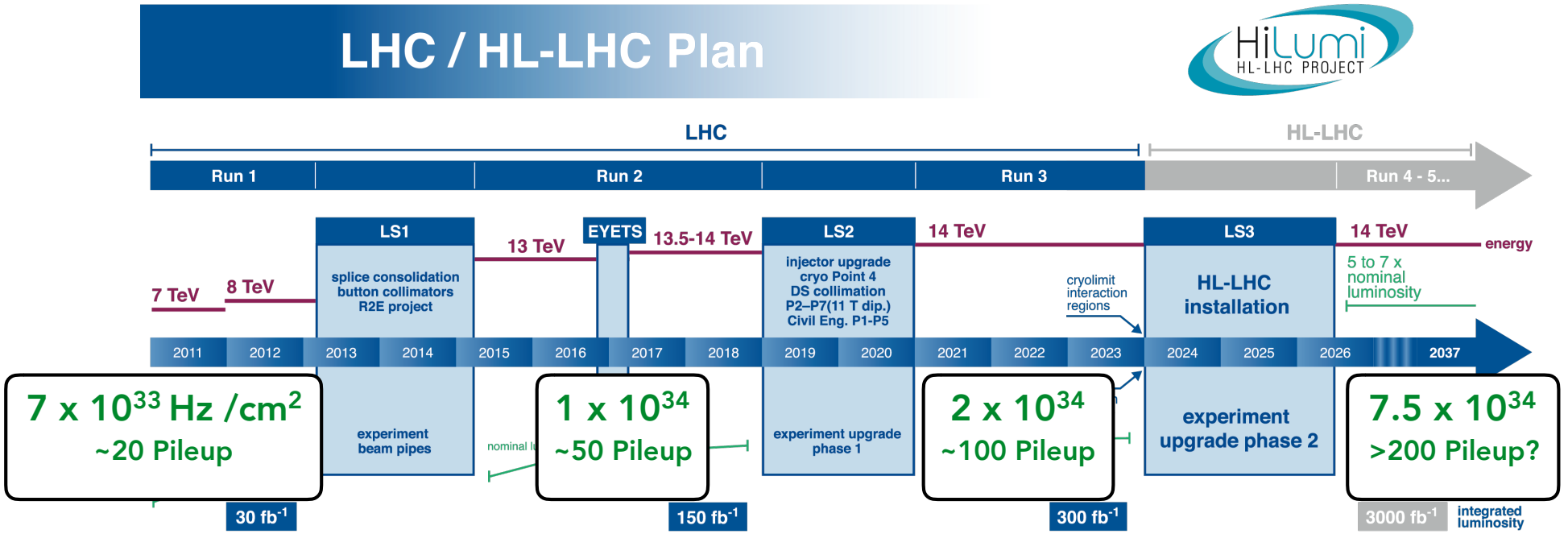




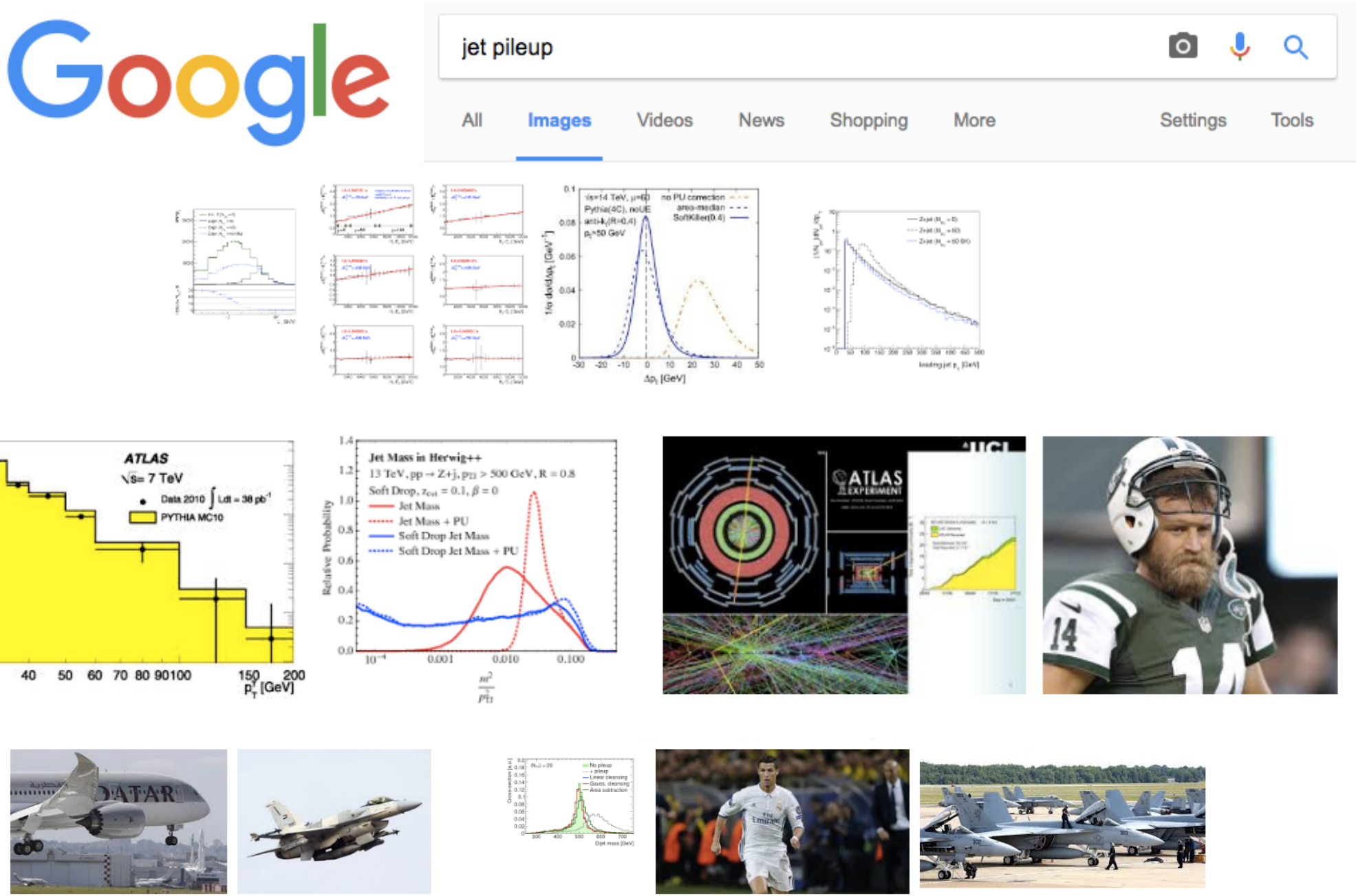
Looking to the Future

Where to go from here?

- Many search results have dramatically increased reach compared to 8 TeV
  - Probing very high masses
- What's going to happen next?
  - More data, of course
  - **Higher instantaneous luminosities**
  - **Higher pileup**
  - Higher energy (?)
- We are already testing ideas to ensure that we can remain efficient in identifying these objects with jet substructure

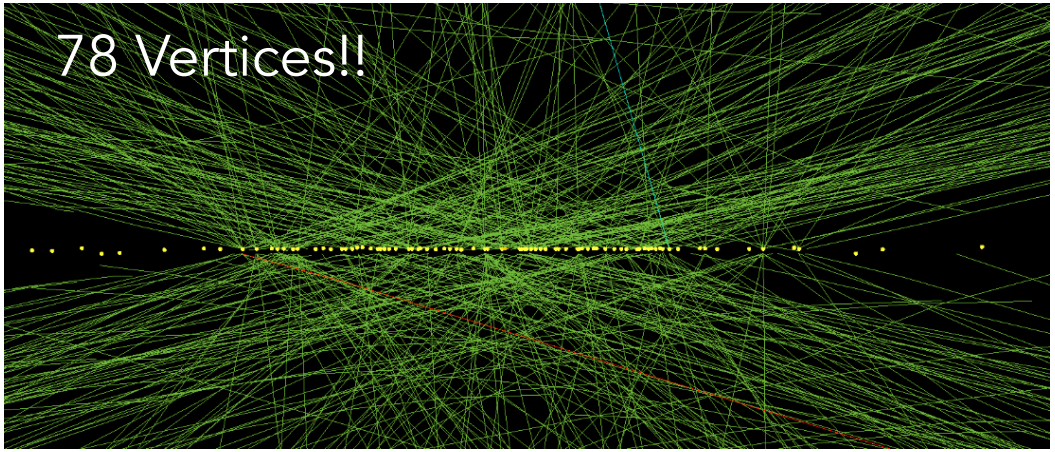


Pileup


Google
jet pileup
All
Images
Videos
News
Shopping
More
Settings
Tools
ATLAS
√s= 7 TeV
Data 2010 ∫ Ldt = 38 pb⁻¹
PYTHIA MC10
Jet Mass in Herwig++
13 TeV, pp → Z+j, p_TJ > 500 GeV, R = 0.8
Soft Drop, z_cut = 0.1, β = 0
Jet Mass
Jet Mass + PU
Soft Drop Jet Mass
Soft Drop Jet Mass + PU
Relative Probability
ATLAS EXPERIMENT
QATAR

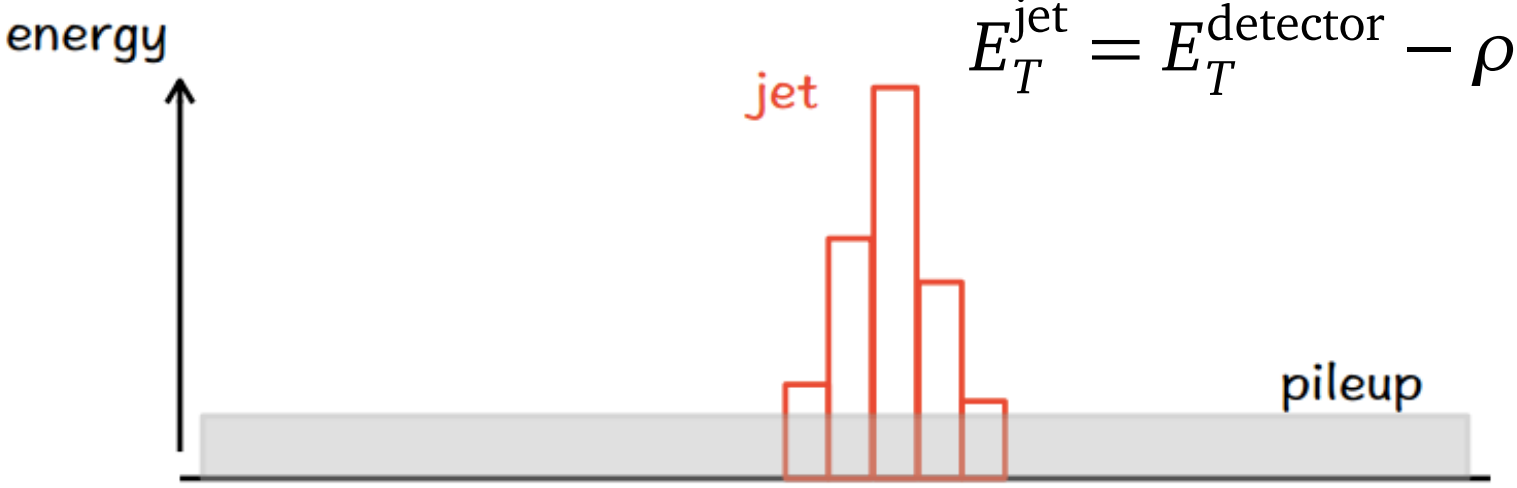
Pileup

- In the future running, we can expect up to 200 pileup events in each bunch crossing of the LHC!



- Can easily determine which charged particles are from pileup vertices
  - Reconstructed tracks point back
  - "Charged Hadron Subtraction" (CHS)
  - What about neutrals??

- Assume uniform energy distribution of pileup energy ρ

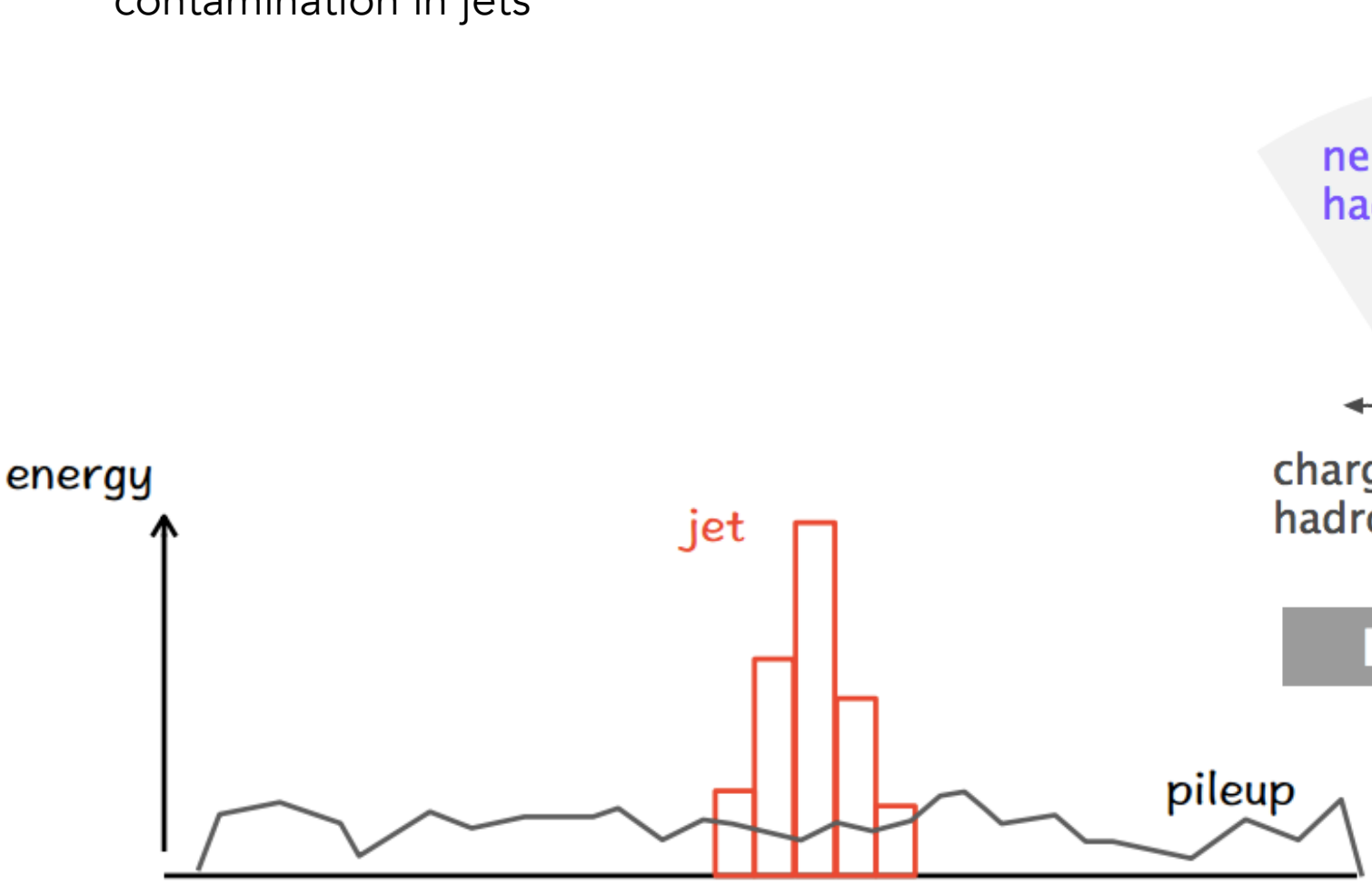


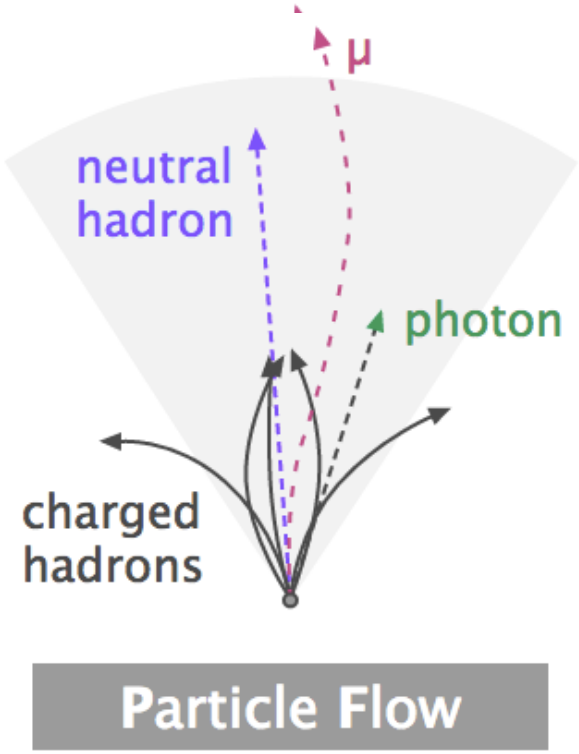
$$E_T^{\text{jet}} = E_T^{\text{detector}} - \rho \cdot A_{\text{jet}}$$

PUPPI arXiv:1407.6013

- The **p**ile**u**p **p**er-**p**article **i**dentification algorithm is designed to use the particle information in the event to improve mitigation of pileup contamination in jets

- We can use tracking information to decide probabilistically if a particle originated from pileup or not



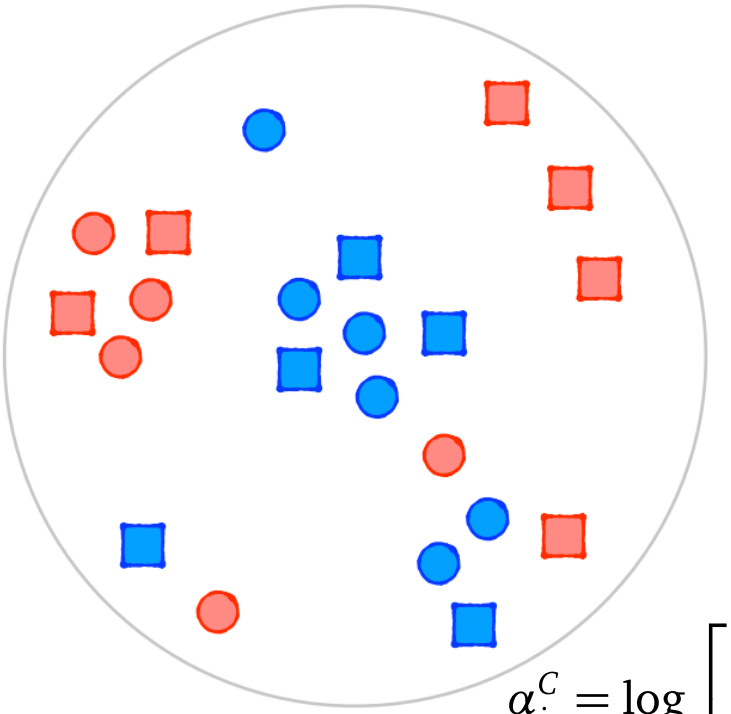


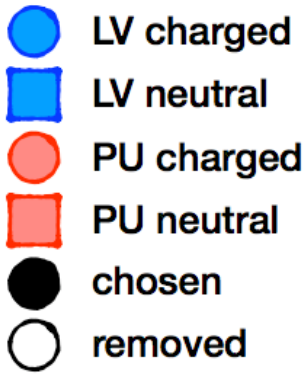
PUPPI

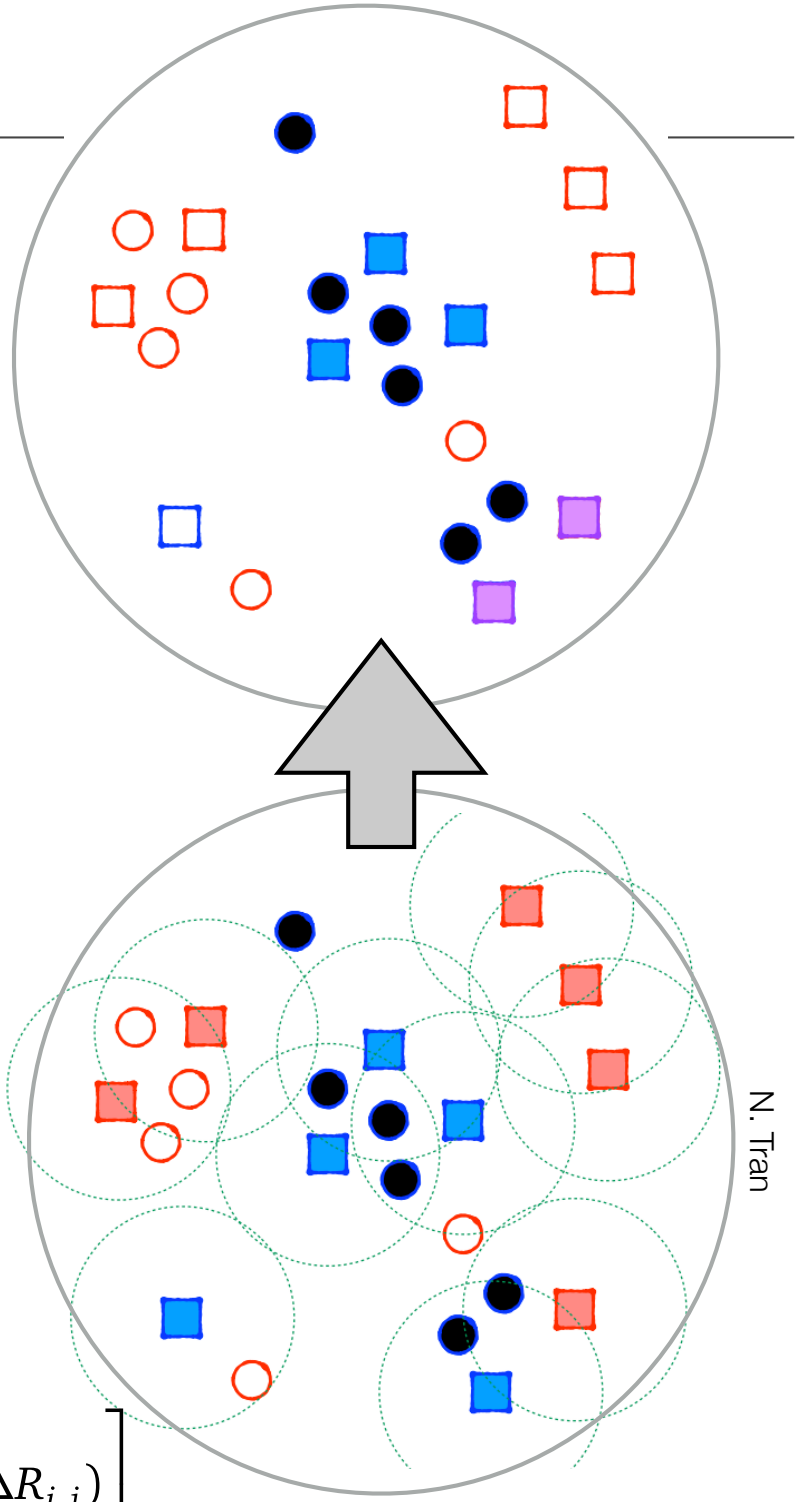
arXiv:1407.6013

- Easy to use tracking information to find which particles originated from the primary vertex

- If neutral particles are mostly surrounded by charged particles from pileup activity, they are probably pileup as well







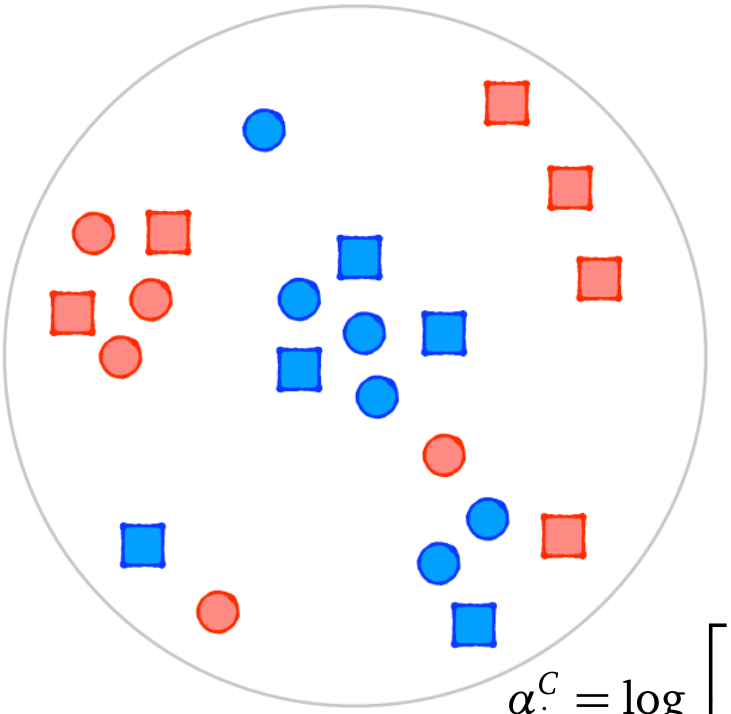
$$\alpha_i^C = \log\left[\sum_{j\in \mathrm{Ch,LV}} \frac{p_{T,j}}{\Delta R_{i,j}}\Theta(R_0 - \Delta R_{i,j})\right]$$

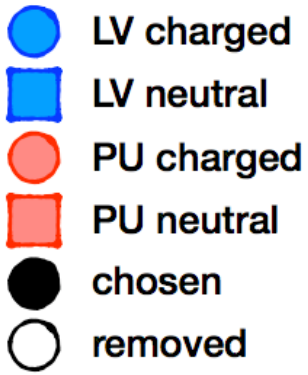
N. Tran

PUPPI

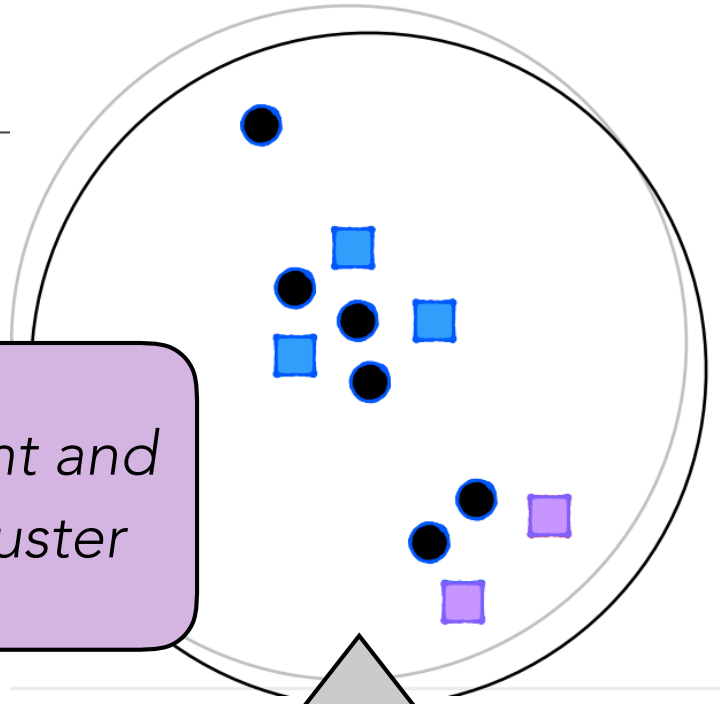
arXiv:1407.6013

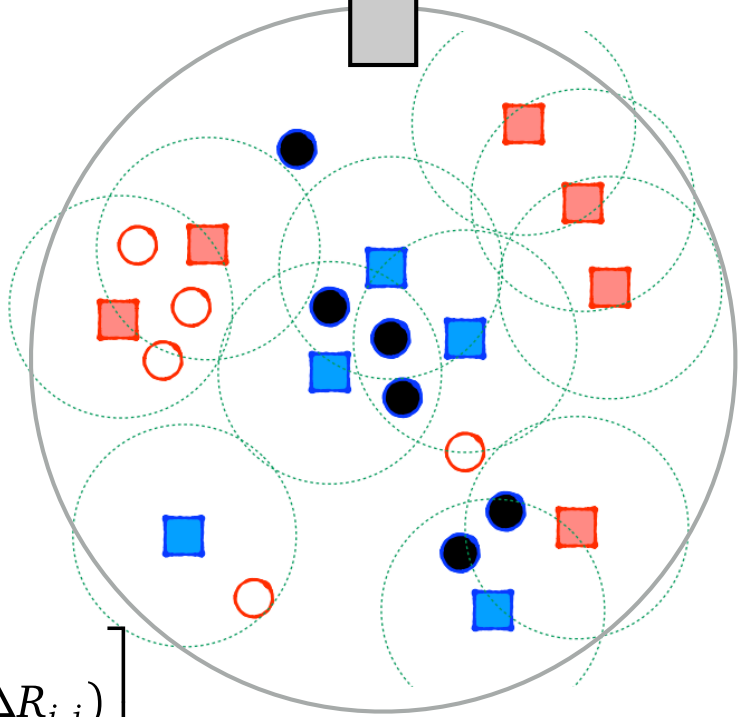
- Easy to use tracking information to find which particles originated from the primary vertex
- If neutral particles are mostly surrounded by charged particles from pileup activity, they are probably pileup as well





Weight and re-cluster



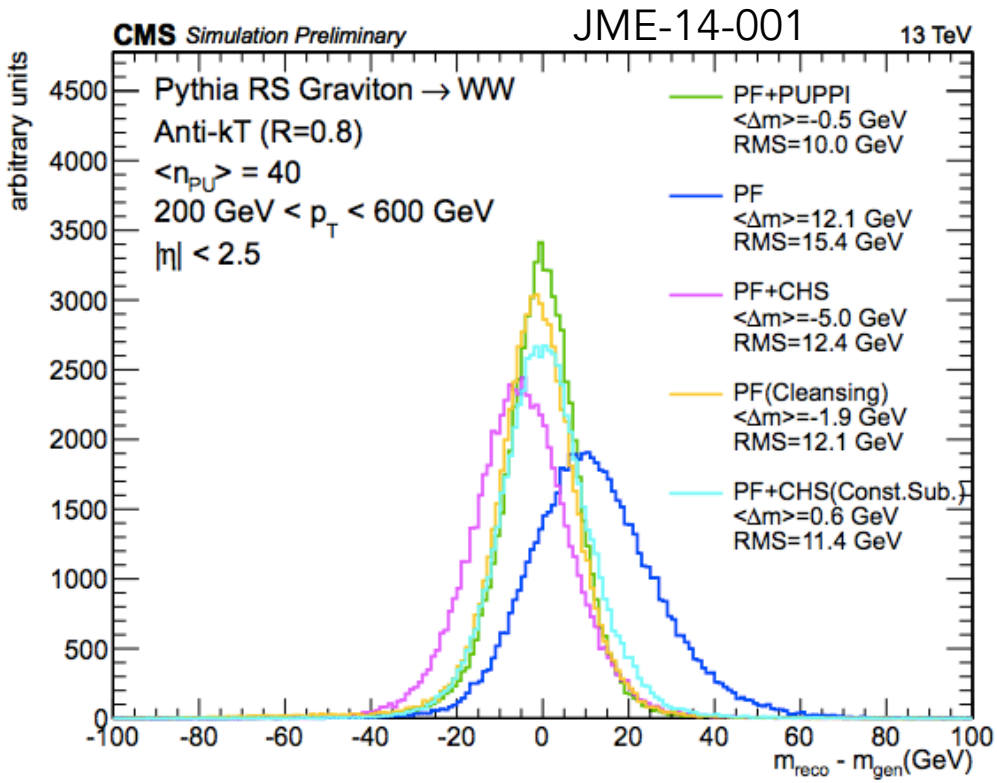


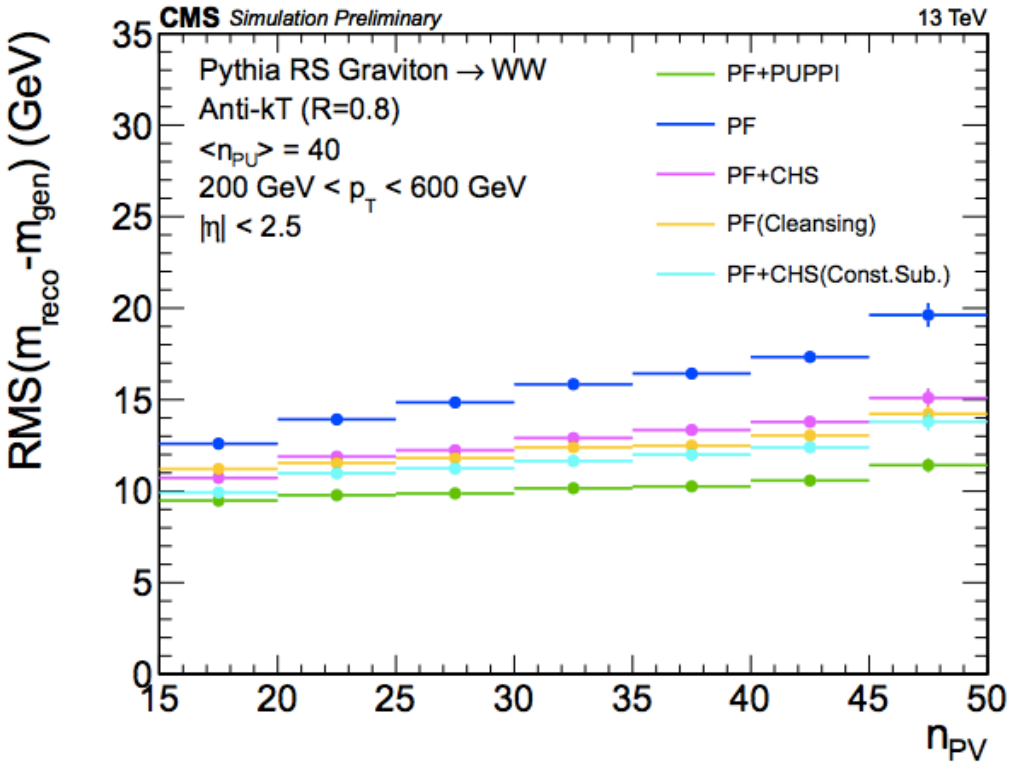
$$\alpha_i^C = \log\left[\sum_{j\in \mathrm{Ch,LV}} \frac{p_{T,j}}{\Delta R_{i,j}} \Theta(R_0 - \Delta R_{i,j})\right]$$

N. Tran

PUPPI Performance

- Improves jet mass resolution significantly
  - Stable as a function of pileup vertices
- Being applied to latest round of CMS searches now
- Also useful for quantities such as MET, lepton isolation

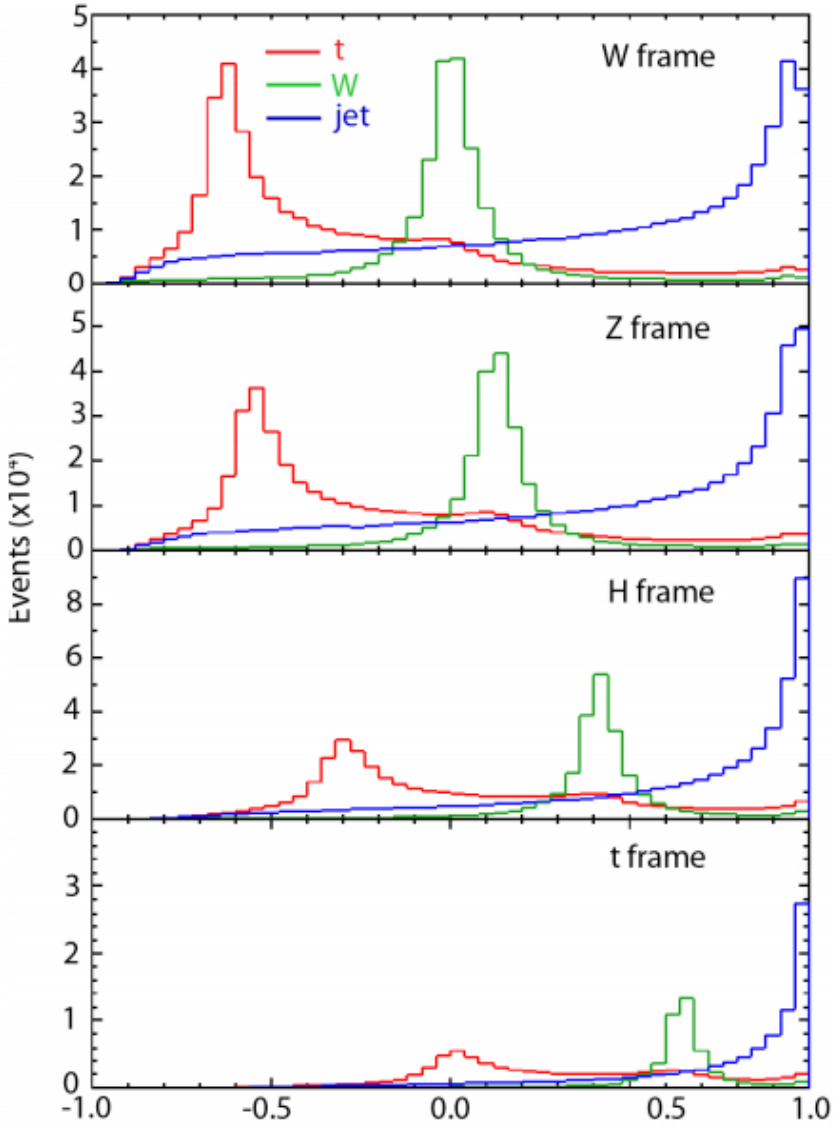


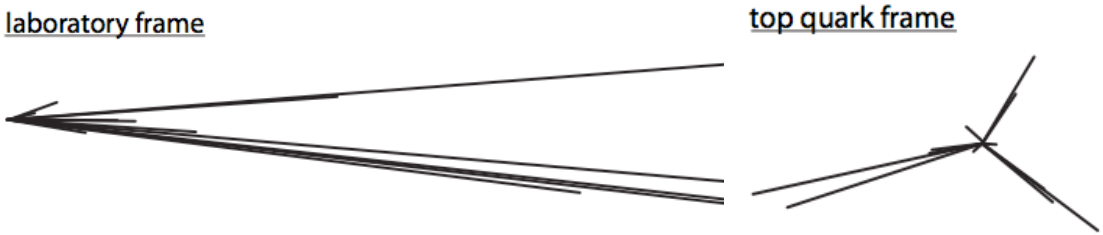


New Substructure Algorithms

arXiv:1606.06859

- Many new substructure algorithms also on the market
- A new method is being explored called the Boosted Event Shape Tagger ("BEST")
- What if we "guess" the particle origin of a jet
  - Extra information can be obtained with different guesses, i.e. W/Z/H/t
- Use boosted reference frames for each hypothesis, compare jet constituent distributions and quantities in each



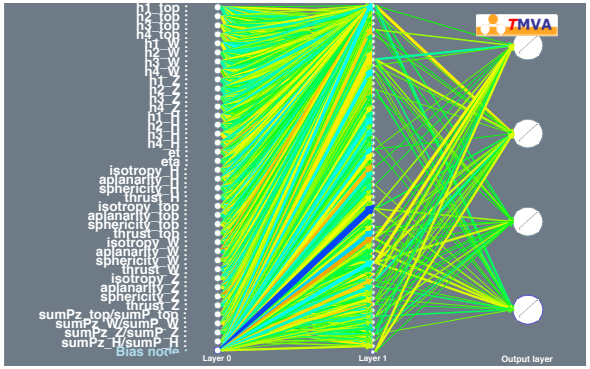


$$A_L = \left[\sum_{\text{jets}} p_z^j\right] / \left[\sum_{\text{jets}} p^j\right]$$

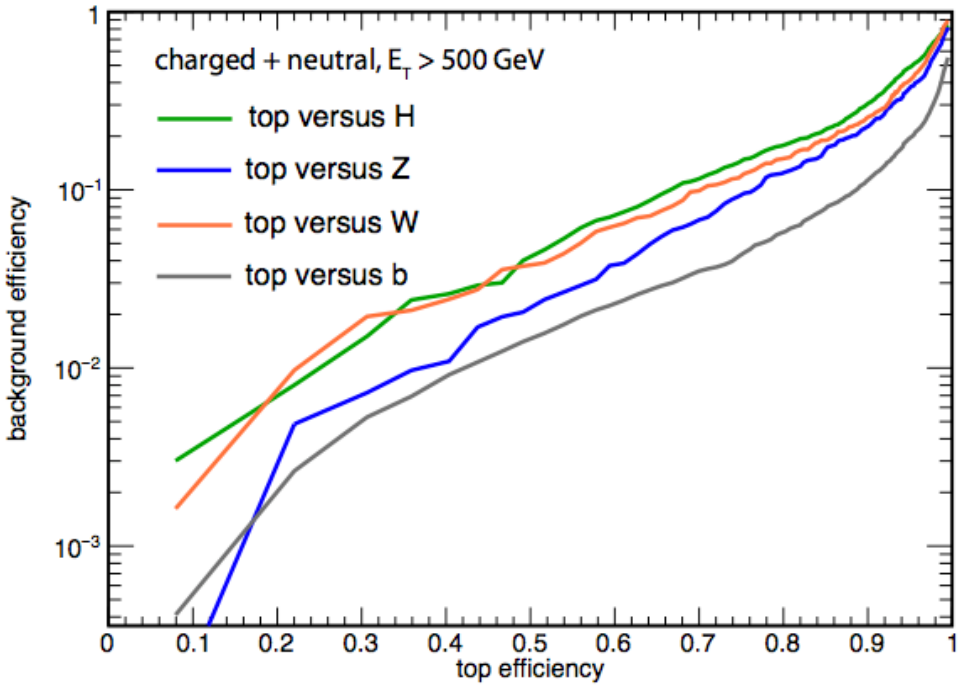
Boosted Event Shape Tagger

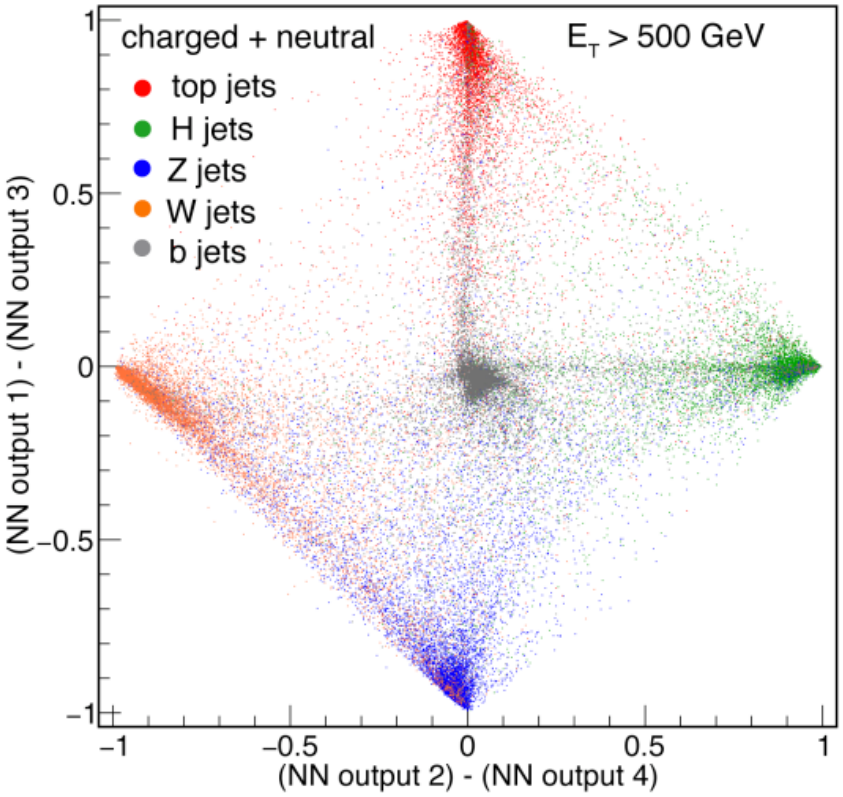
arXiv:1606.06859

- Use these angular quantities and distributions to train a NN to separate the different particle species
  - One output for each particle species



- Discrimination can be obtained for each of the particle species
  - ~70% signal efficiency with ~95% background rejection





- Currently being optimized with CMS data for use in analyses!

Conclusions

Conclusions

- CMS has well-developed tools to take advantage of **jet substructure**
  - **Enhances discovery potential in high-p_T regimes**
  - Tools are in use in many analyses
- It is critical to understand these algorithms in the case of a hint of signal events
- We continue to work on optimization and development of algorithms for the changing running conditions
  - **Planning for the next discovery!**
- **Thank you for your attention!**
  - **Comments, questions?**