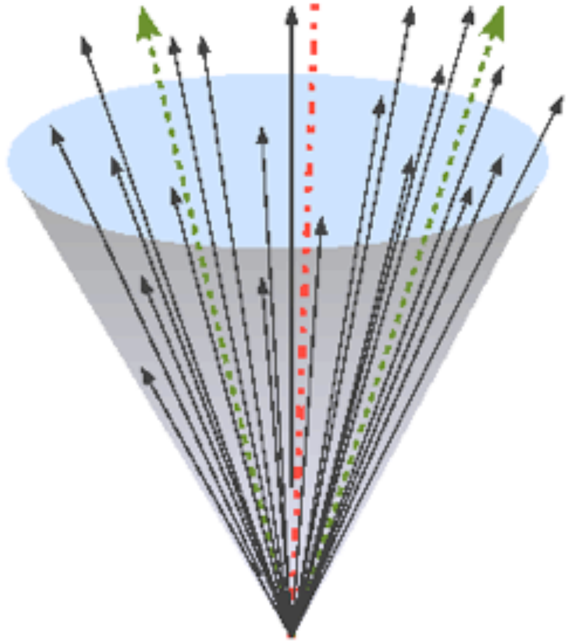
A case study of using substructure analysis in new physics searches

Kalanand Mishra
Fermilab



QCD Tools for LHC Physics
November 14, 2013

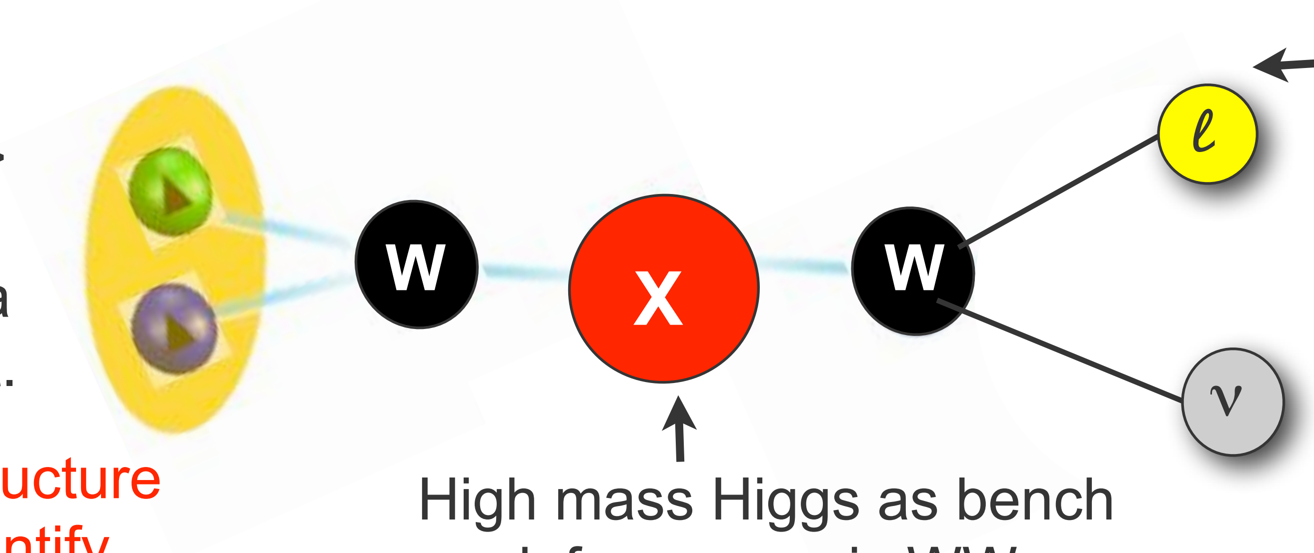
The case study ....

Search for high mass WV resonances (V = W or Z)

http://cds.cern.ch/record/1546778

CMS HIG-13-008

For W p_T > 200 GeV, generally a merged jet.



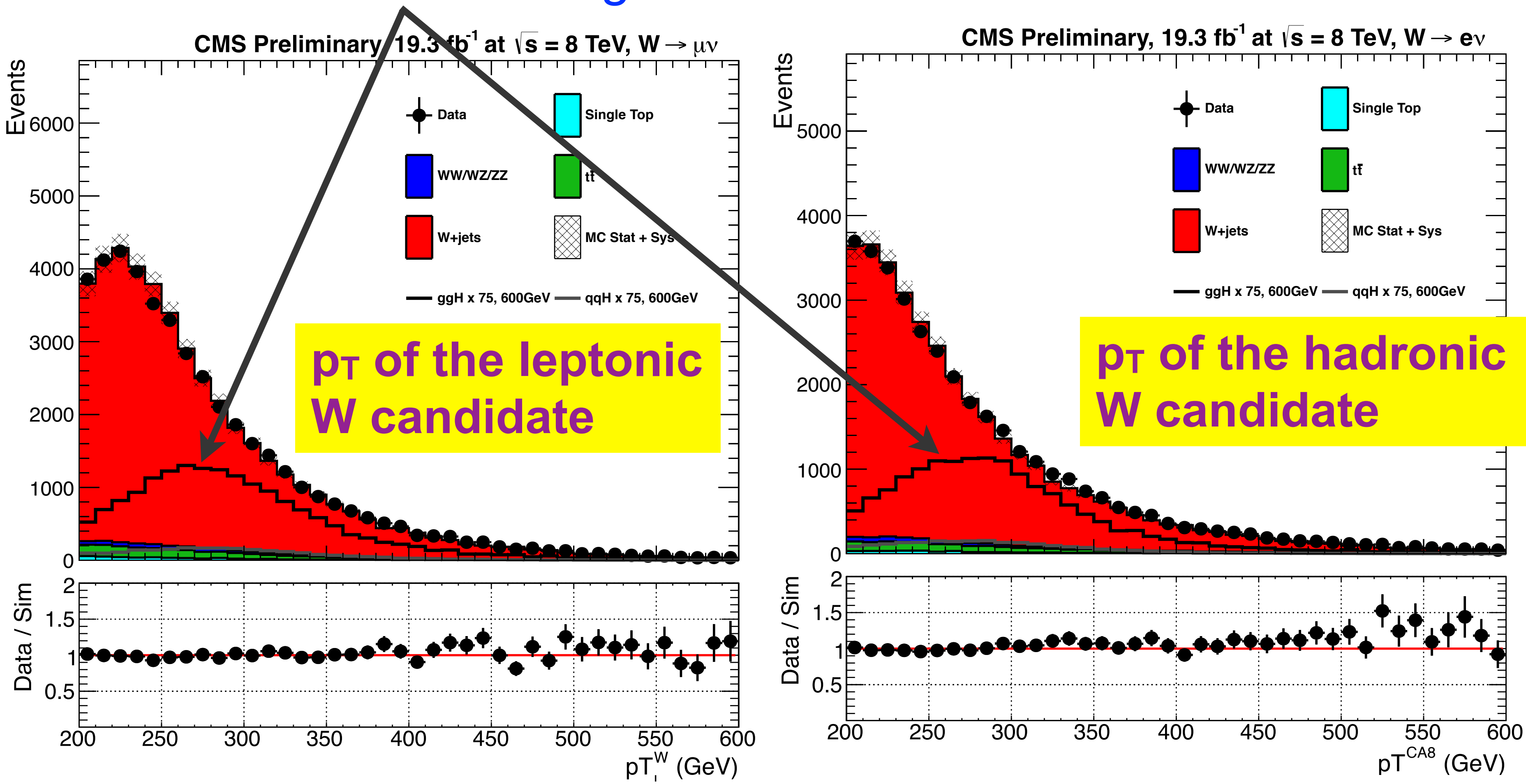
Used to trigger events

Use substructure tools to identify hadronic W (or Z)

High mass Higgs as bench mark for a generic WW resonance of M > 600 GeV

Why the boosted analysis?

Because the signal events have boosted Ws



Also, otherwise the background becomes overwhelming

A quick summary of the event selection

Both leptonic and hadronic W p_T > 200 GeV

- Exactly one lepton with p_T > 30/35 GeV for μ/e
- MET > 50/70 GeV
- At least one CA8 jet with p_T > 200 GeV
- Veto presence of any b-tagged AK5 jets in the event

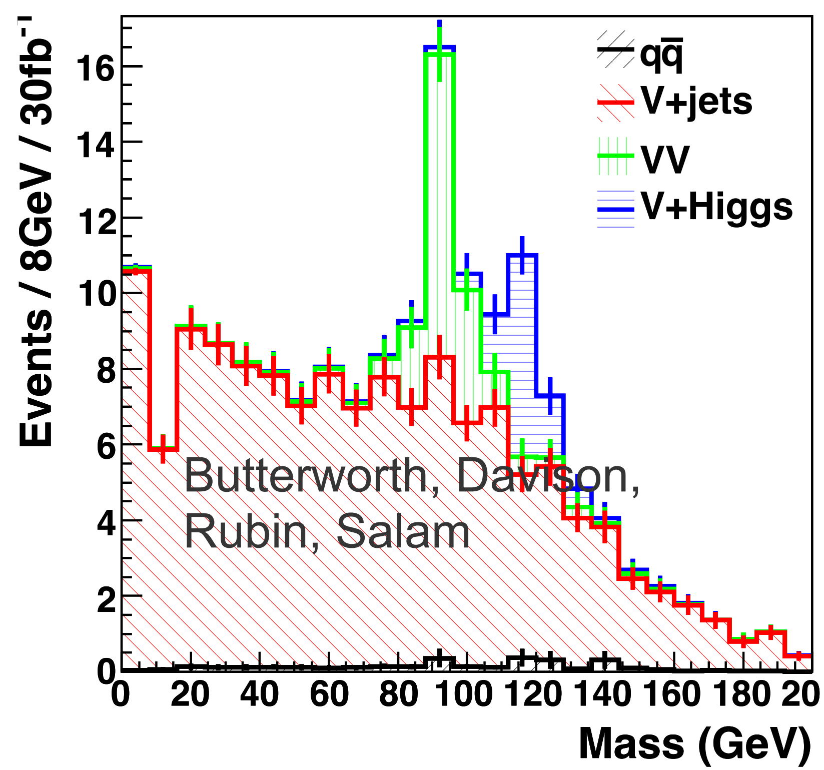
Additional cleaning cuts

- $\Delta R_{l,j} > \pi/2$; distance between the lepton and the jet
- $\Delta\Phi_{\not{E}_T,j} > 2.0$; azimuthal separation between jet & MET
- $\Delta\Phi_{V,j} > 2.0$; azimuthal separation between the two Ws

I will discuss some aspects of this selection at the end

Use of jet substructure

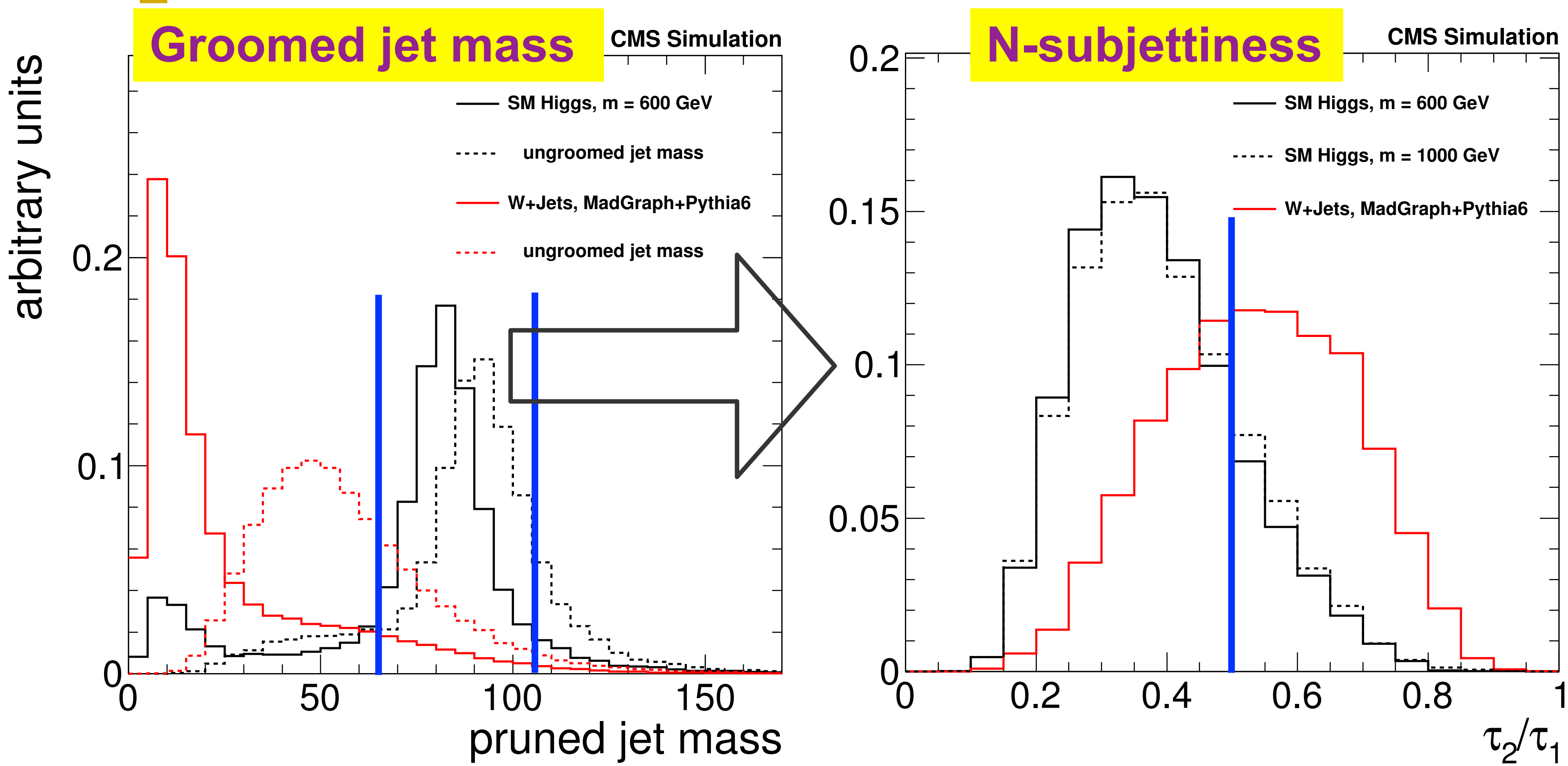
- Jet substructure can be used to reconstruct hadronic decays of boosted "heavy particles": Higgs, W/Z, top
- Several measurements available now: see talk by Jeremy Love



arXiv: 0802.2470

Have developed tools for W (and Z) tagging for the present analysis

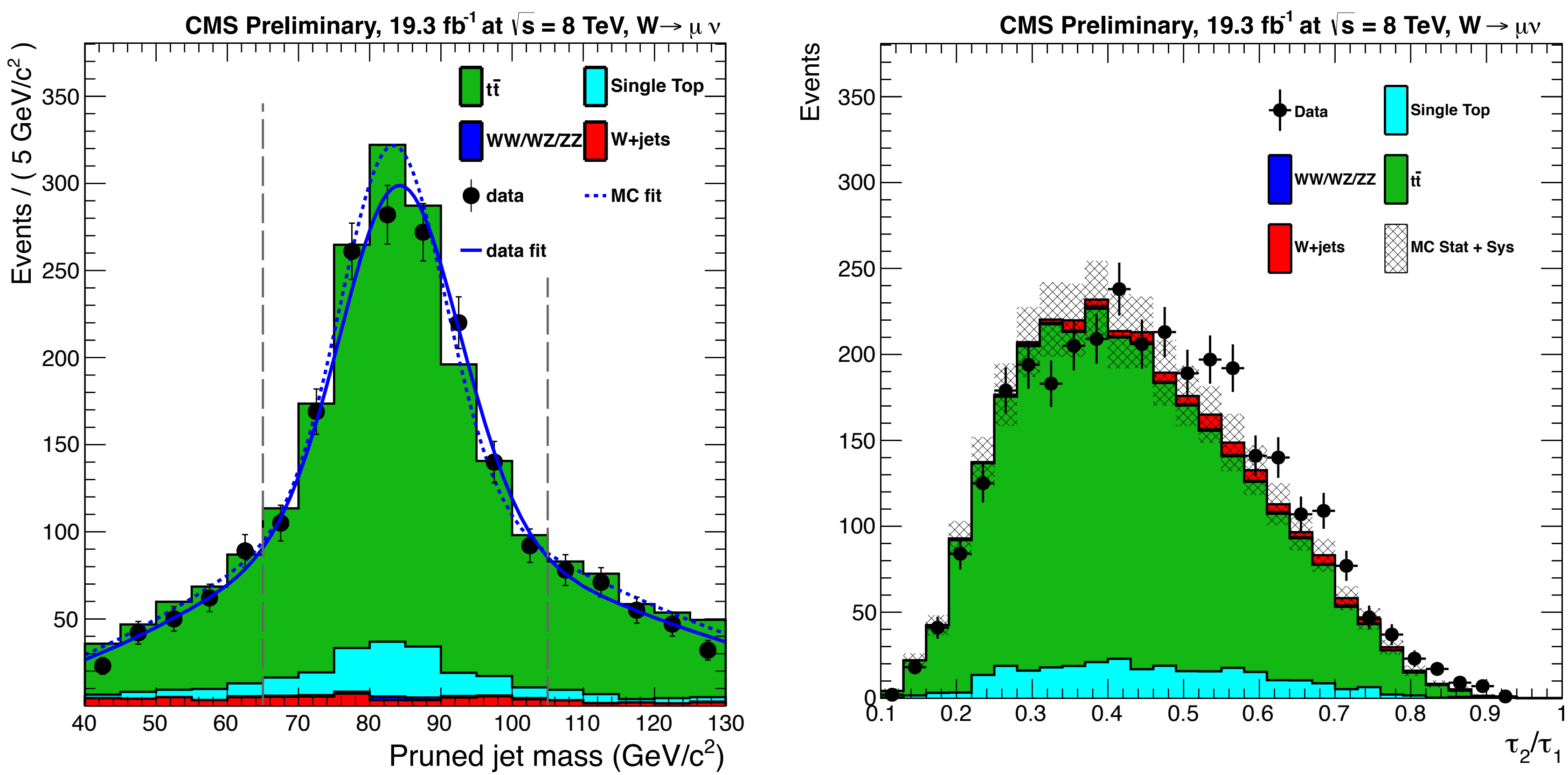
The tools used in this analysis are ...



These two essentially gave the max discrimination. Tried additional observables, e.g., mass drop & MVA, but w/o much gain. Why?

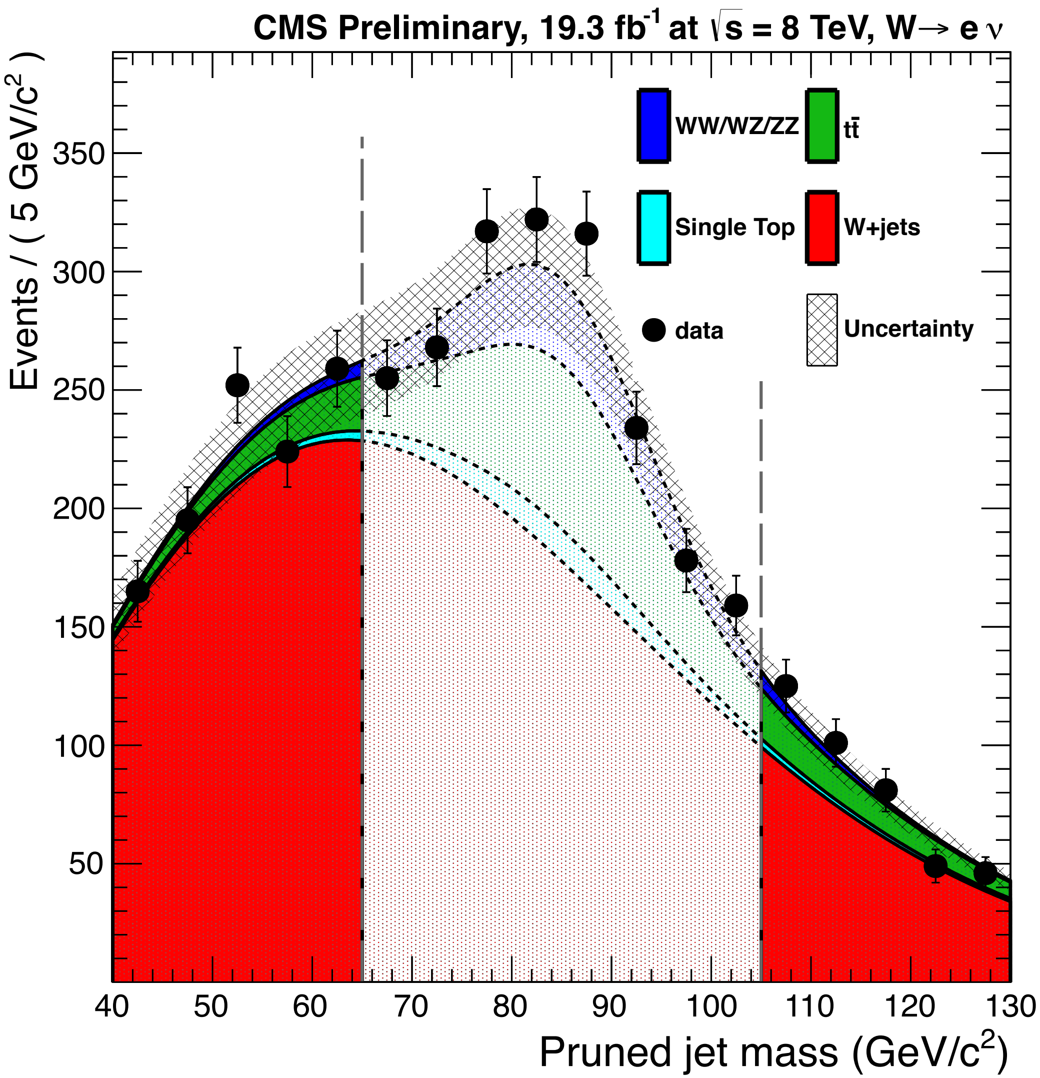
How do we know if all this works in data?

Look in the top control region by requiring ≥1 b-tag



Worse resolution in data. Also find that for W tagging data/MC scale factor = 0.95±0.10 (0.86±0.10) in the μ(e) channel

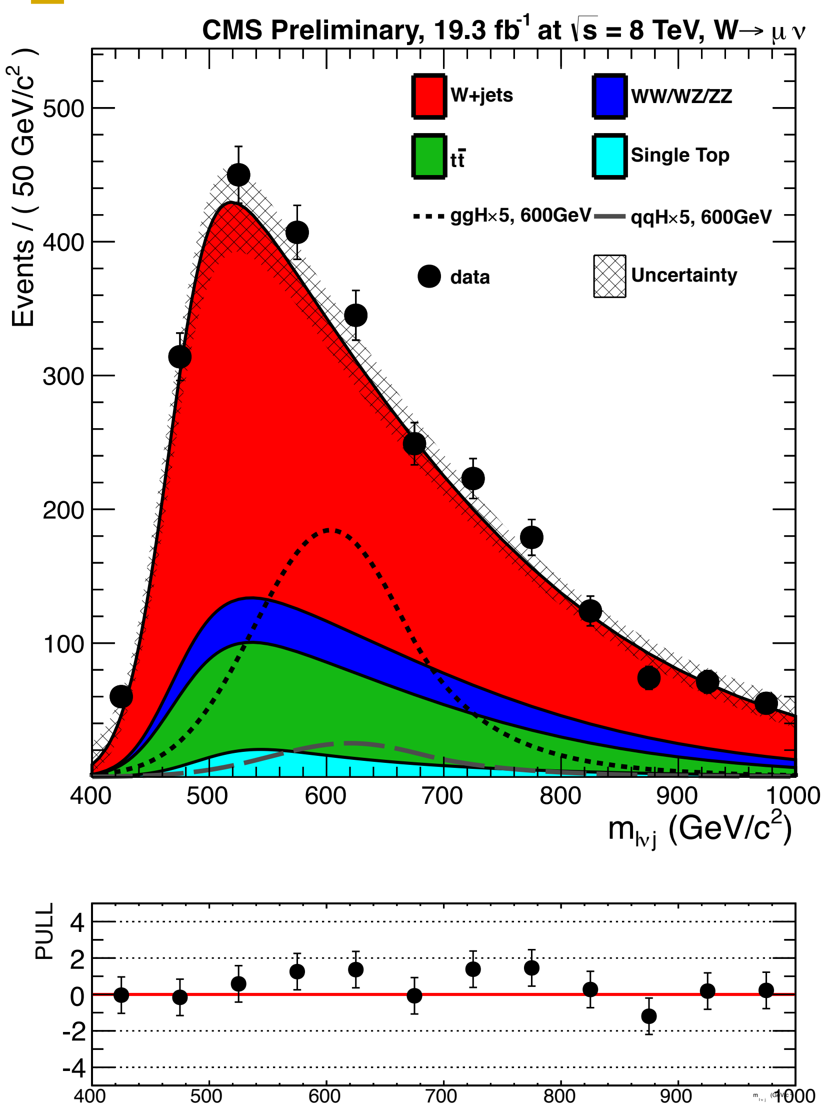
Data sideband can help to estimate background



- Perform a template fit to the data SB to estimate contribution of W+jets
- Top background (ttbar + single top) estimated using b-tagged CR in data
- Other backgrounds taken directly from simulation

All shapes parametrized using guidance from MC

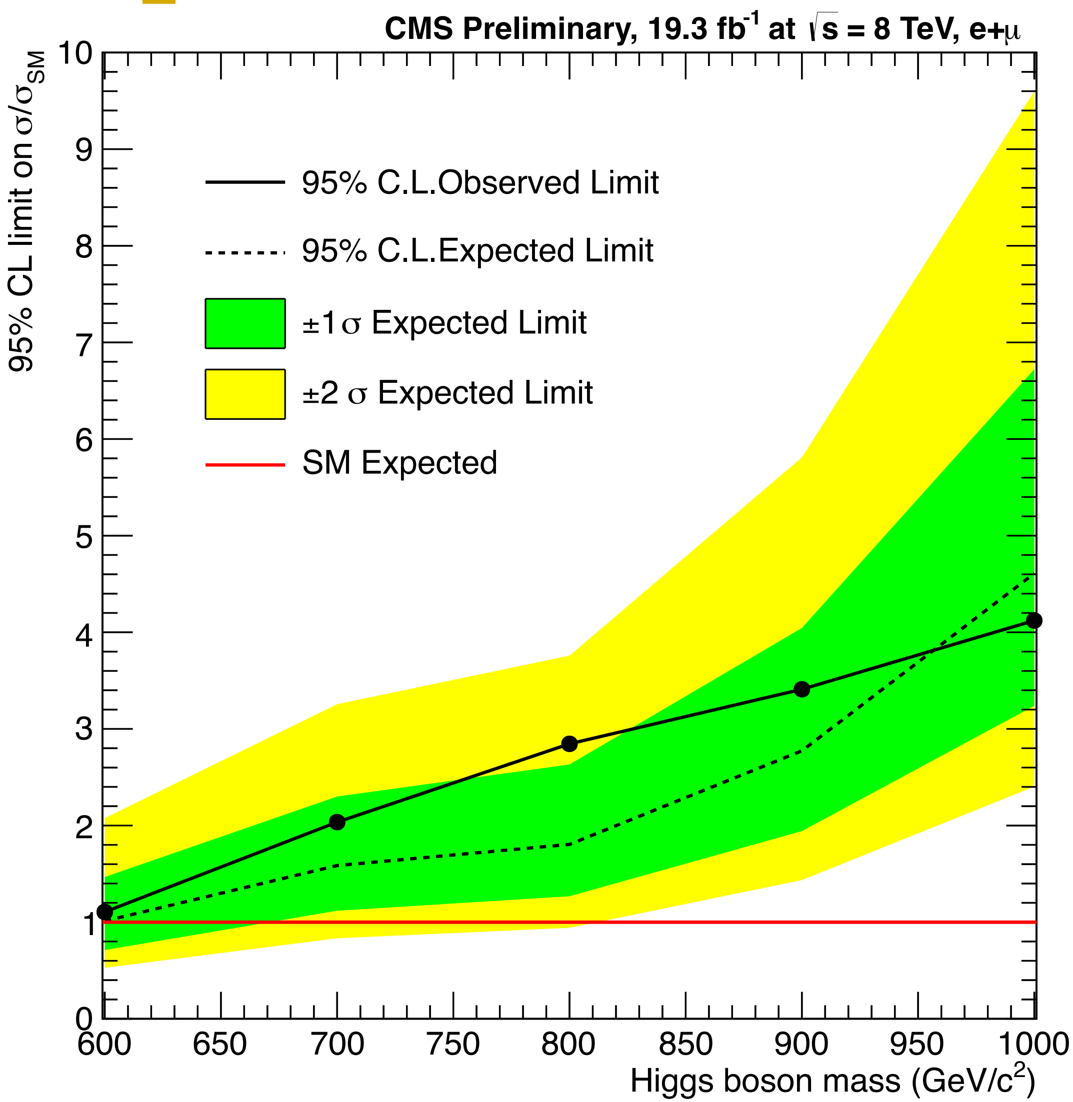
Finally, look at the WW invariant mass



- Normalizations from the previous slide
- W+jets shape taken from data using events in the sidebands in m_J
- All other shapes taken from MC (with smoothing/ parametrization)

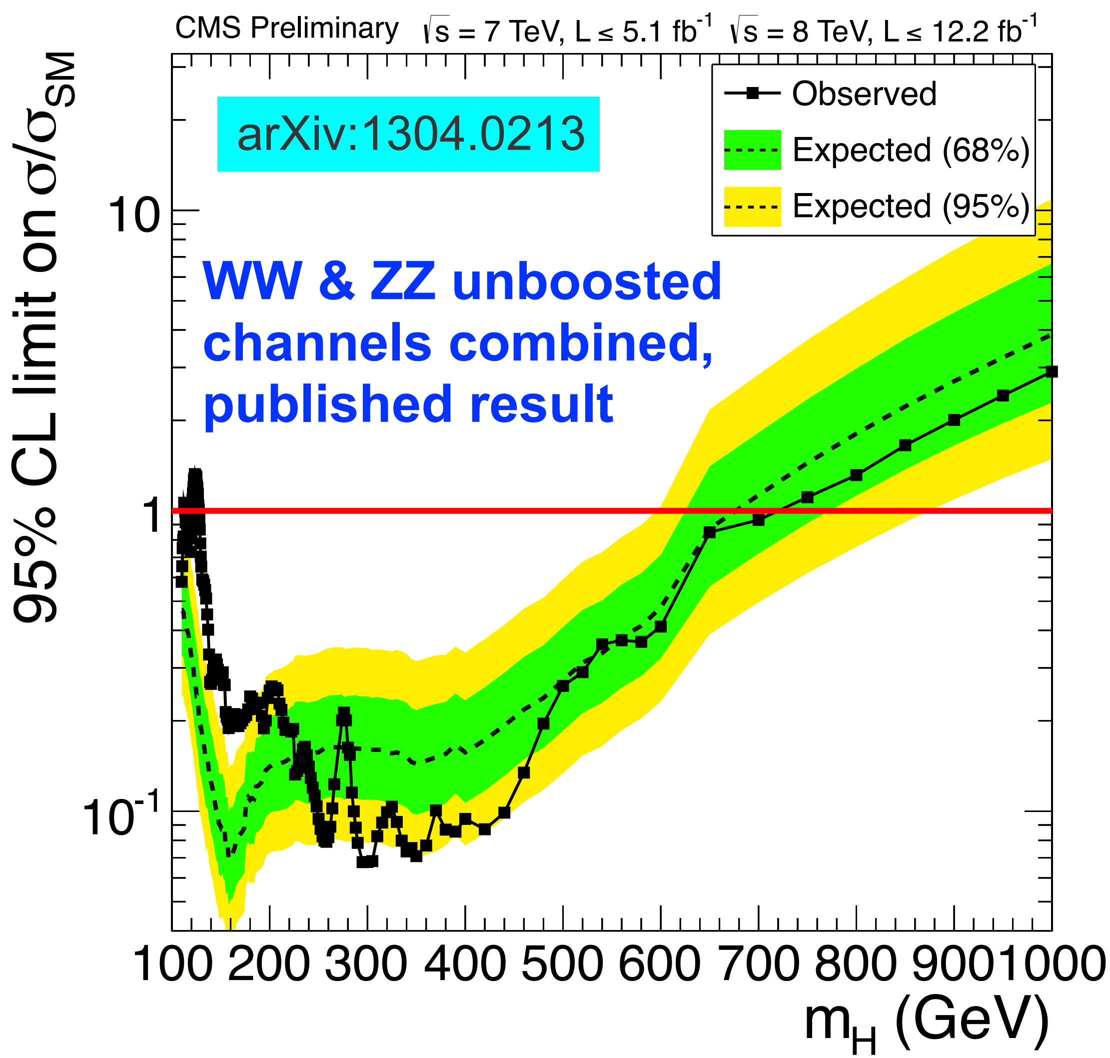
No excess seen, set limits

A high mass Higgs state with SM couplings



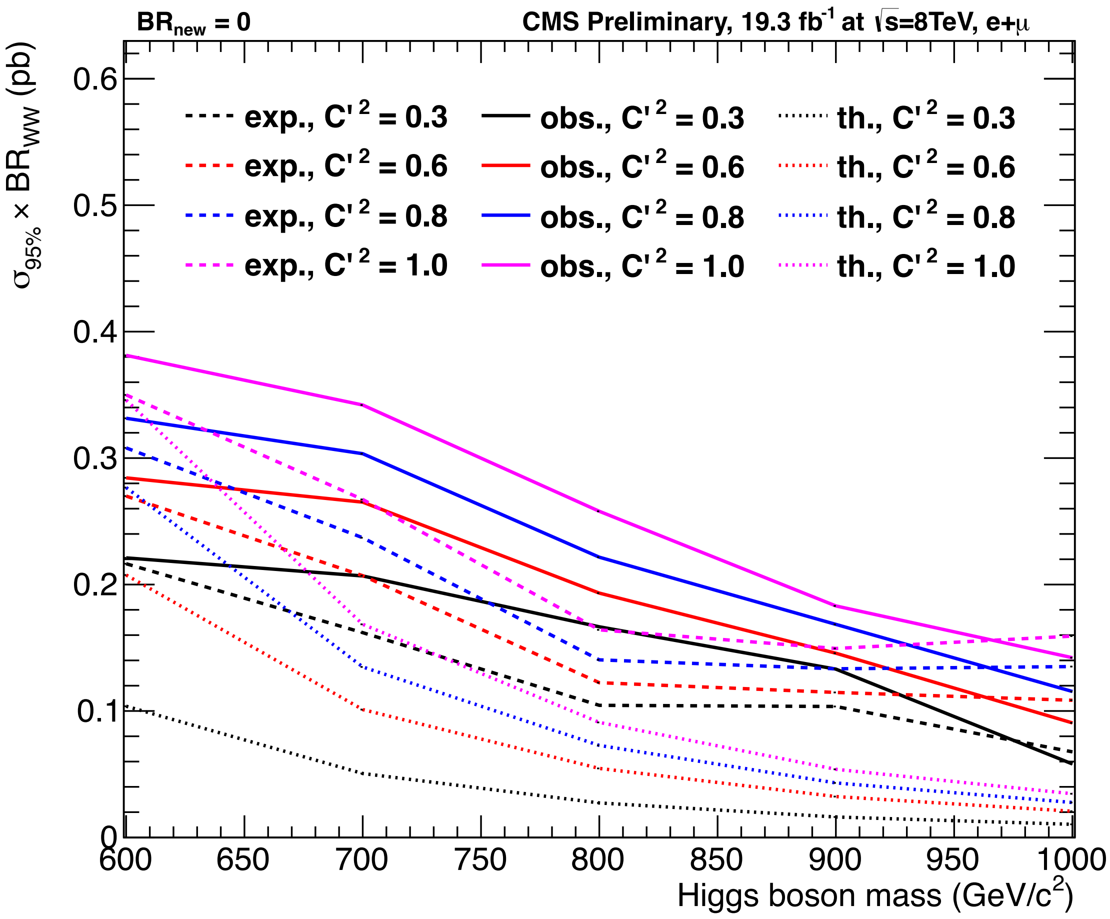
- Use high mass Higgs state with SM couplings as the benchmark signal for WW resonances
- The idea of an SM-like Higgs gets fuzzy at high mass b/c its width ~ m_H
  - -So, also present the result in terms of a narrow particle with modified couplings

BTW, combination of all unboosted channels



- No heavy Higgs or VV resonance up to 800 GeV
- The boosted WW channel alone more sensitive than the combined unboosted result above 800 GeV

A more generic BSM interpretation



In the BSM interpretation, search for an electroweak singlet scalar, where a heavy Higgs boson mixes with H(126)

$$C^2 + C'^2 = 1$$

The heavy Higgs cross cross-section and width are modified as

$$\mu' = C'^2(1 - \mathcal{BR}_{new})$$

$$\Gamma' = \Gamma_{SM} \times \frac{C'^2}{(1 - \mathcal{BR}_{new})}$$

Possible improvements and other issues for future

- In the current iteration used b-tagging on AK5 jets
  - Well-supported & understood, but not a happy situation
  - Possible to (anti) b-tag the CA8 subjets directly
  - Requires investments: operating points (loose, tight, ..), efficiency & fake-rate, ... Some effort underway.

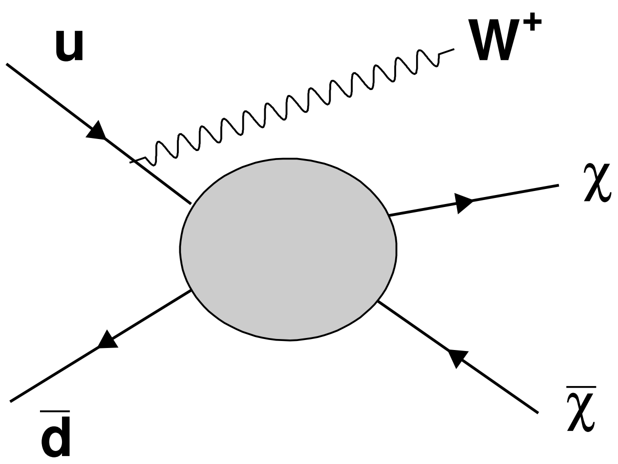
- In the very high p_T regime (W p_T > 500 GeV)
  - Bkgd from (semi) boosted top not very constrained
  - Can veto a 2nd CA8 jet of p_T > 100 GeV, but ....?
  - Other ideas to reject boosted top, “anti-tagger”! ?

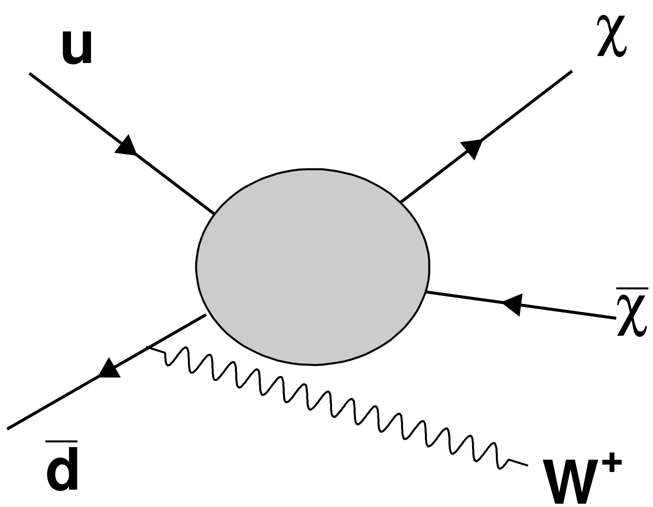
- High efficiency quark-gluon tagging for subjets ?
  - Main challenge is to define control & calibration data

- Most of these manageable. Other ideas, feedback, ... ?

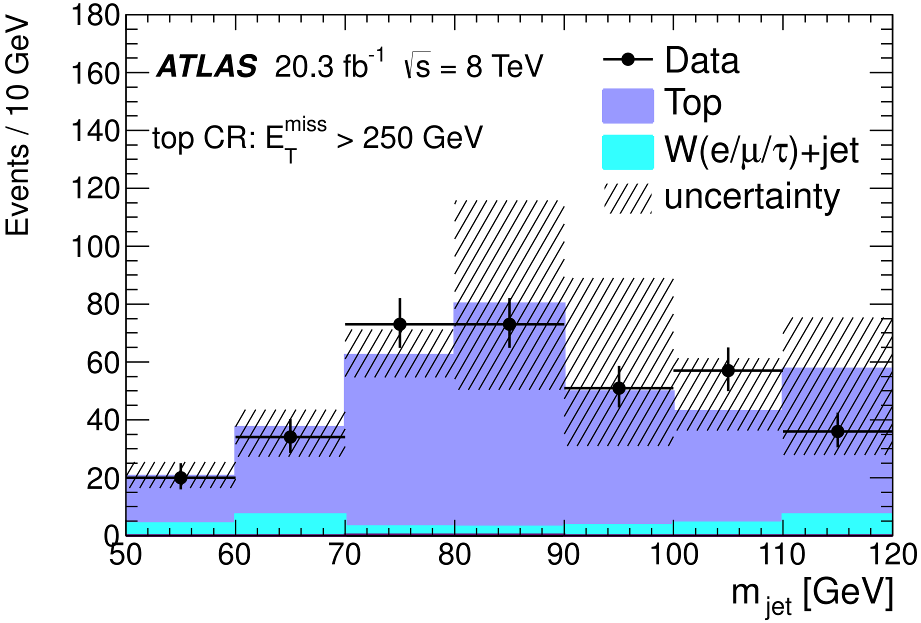
BACKUP SLIDES

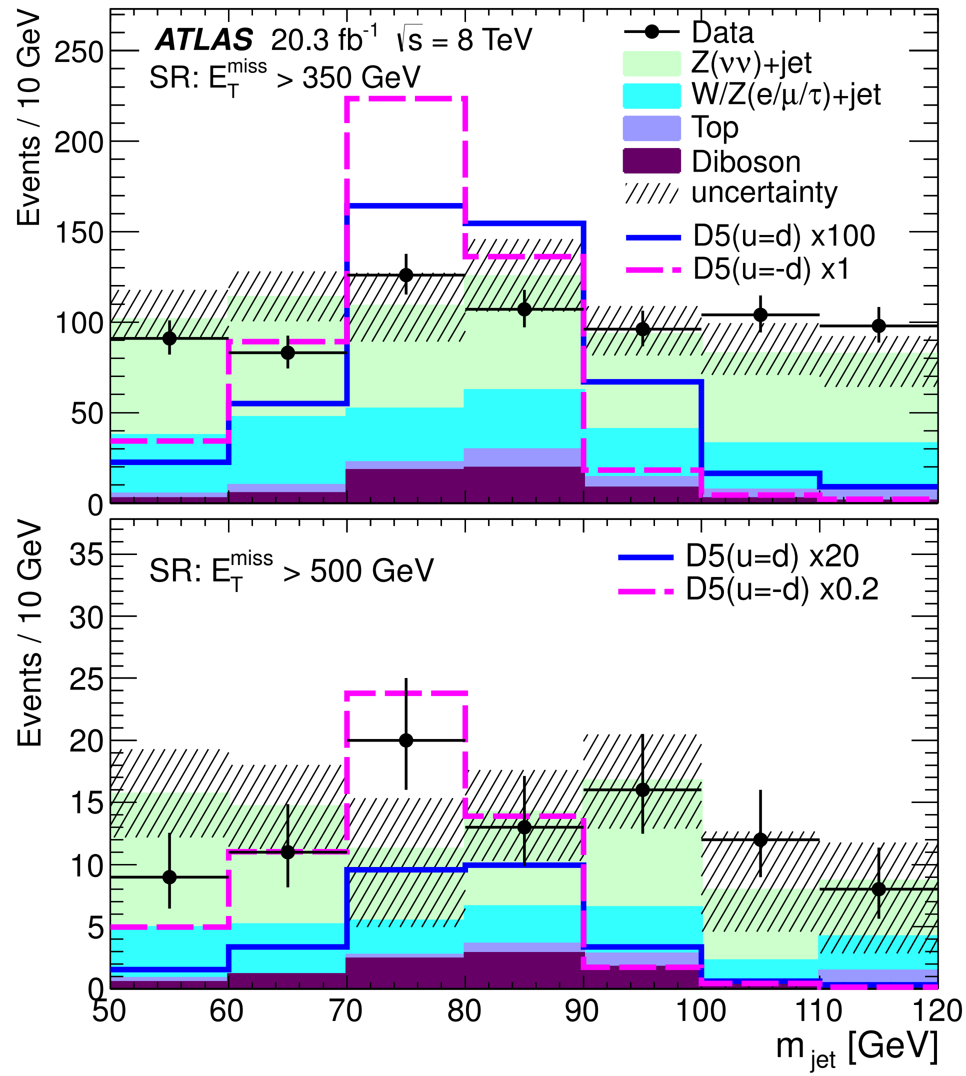
A variation of the above analysis





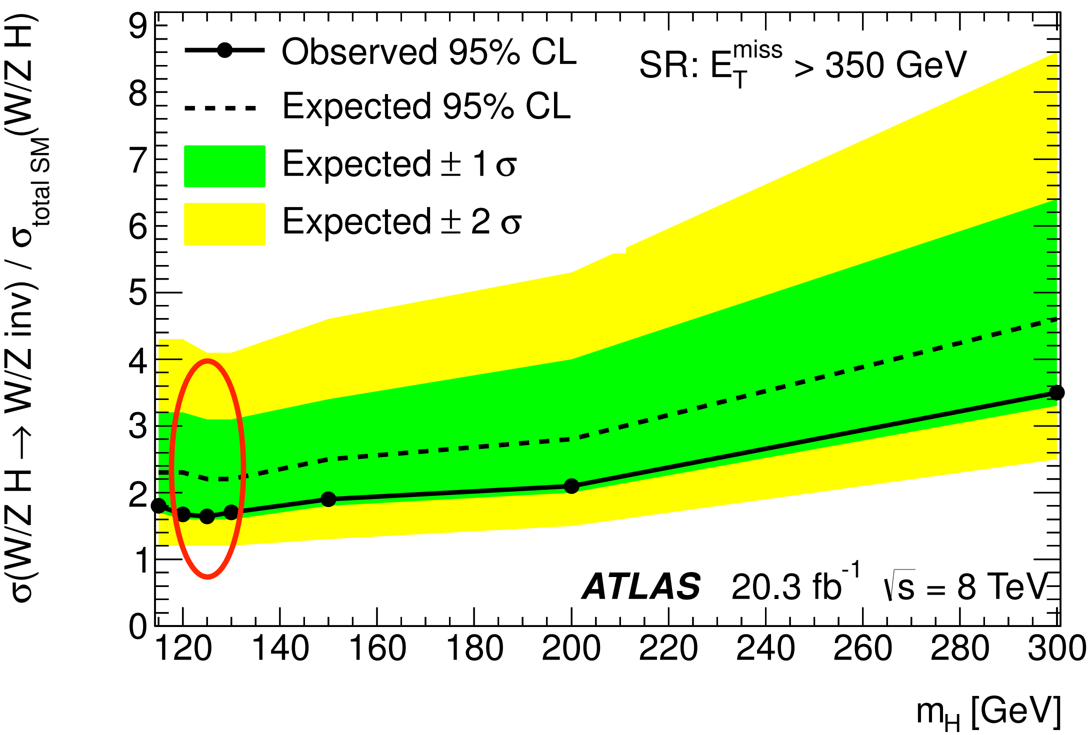
Again, m_J in the top CR and signal region

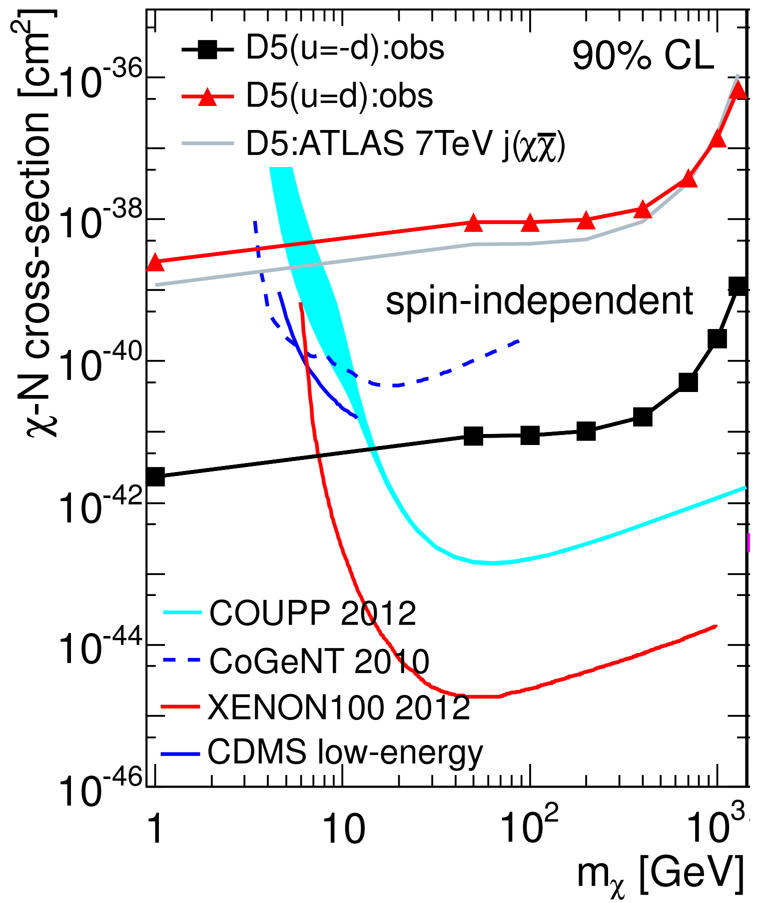




Probing Higgs invisible decay & dark matter

arXiv:1309.4017





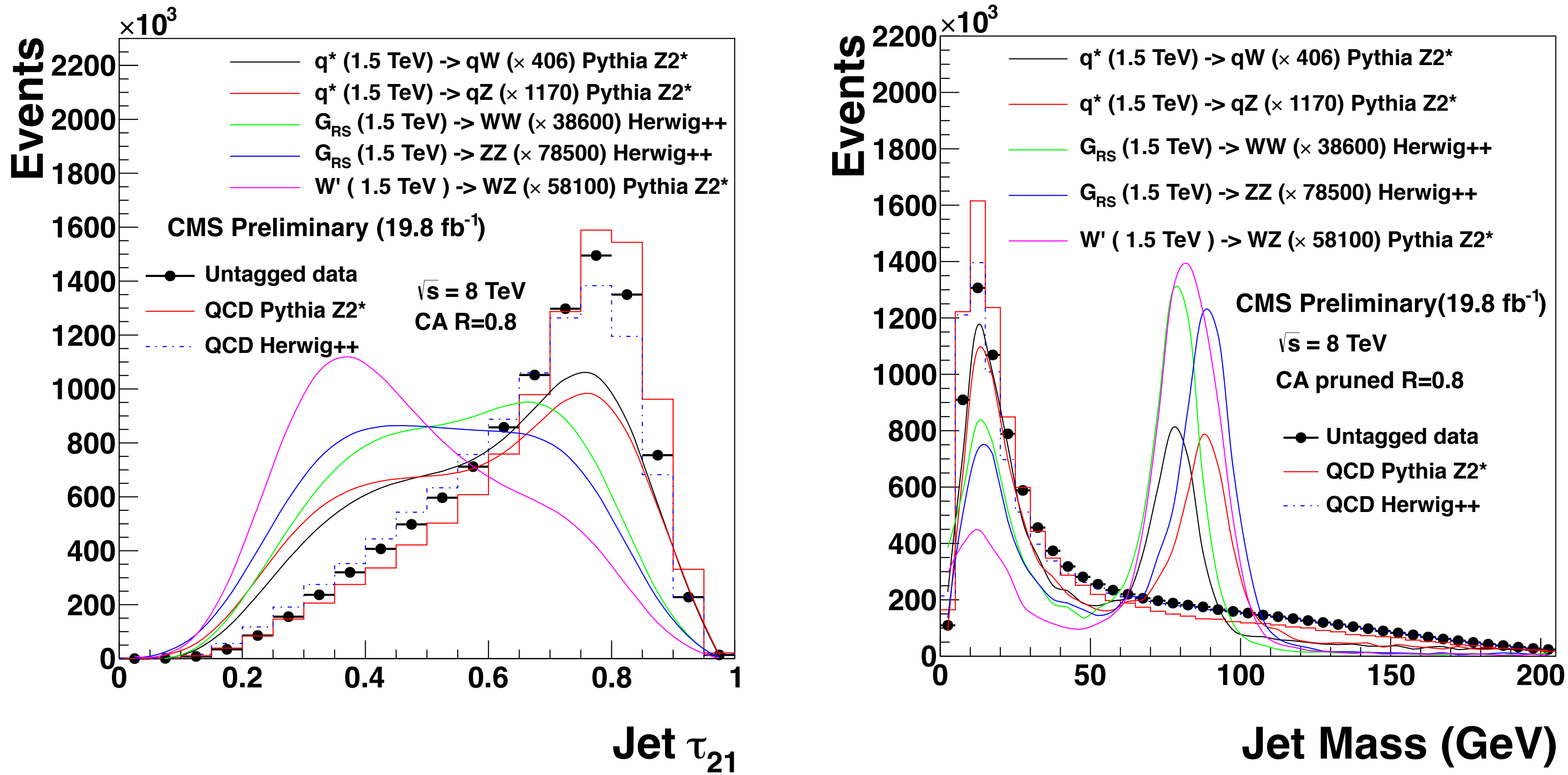
Started probing the Higgs invisible decay “directly” in the associated production mode.

The best limit on WIMP DM mass 1−5 GeV. Comparable to mono-jet limits.

Just for completeness ....

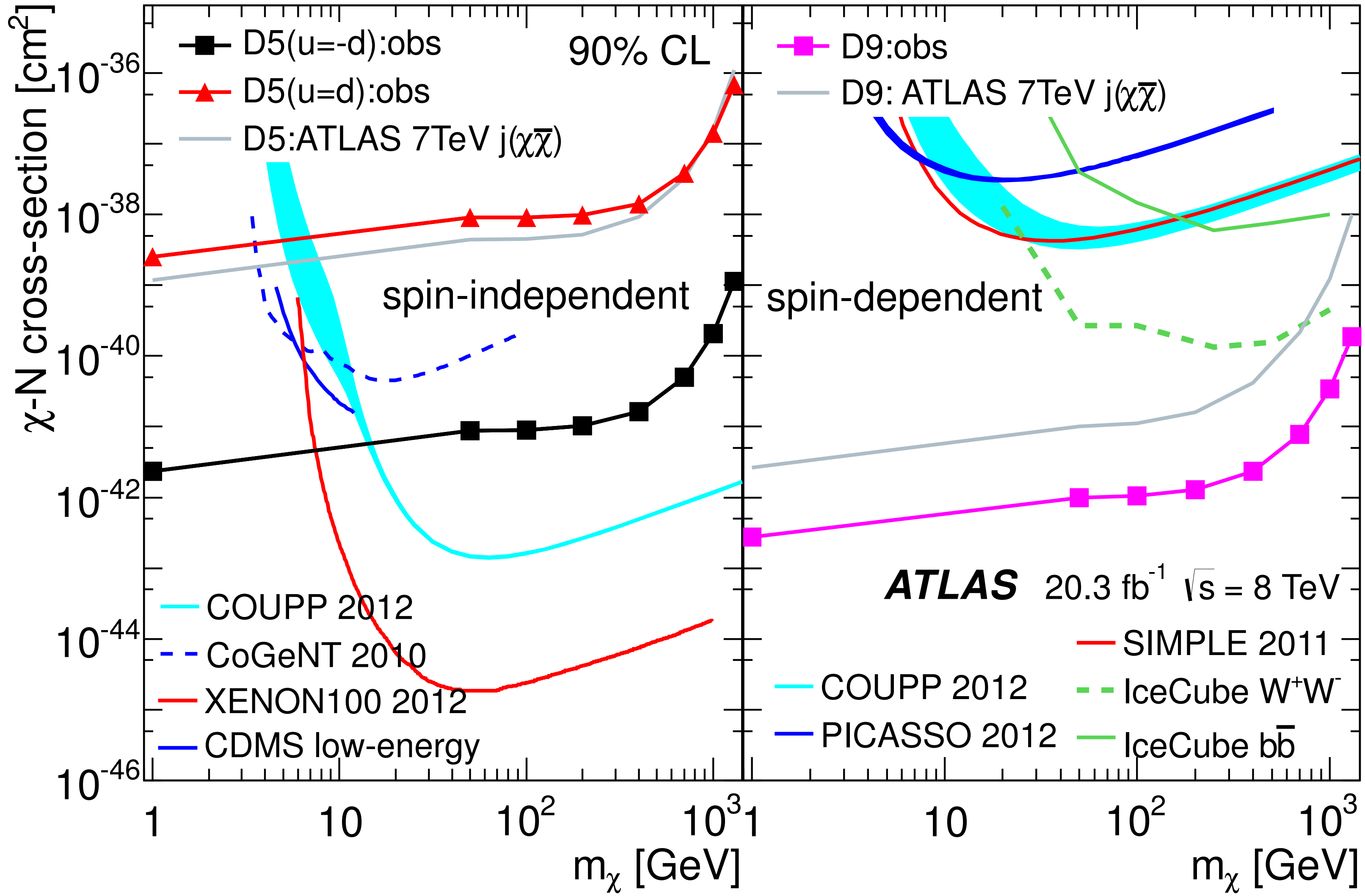
CMS EXO-12-024

Events with 2 boosted jets. Hope to reco WW, WZ, ZZ.

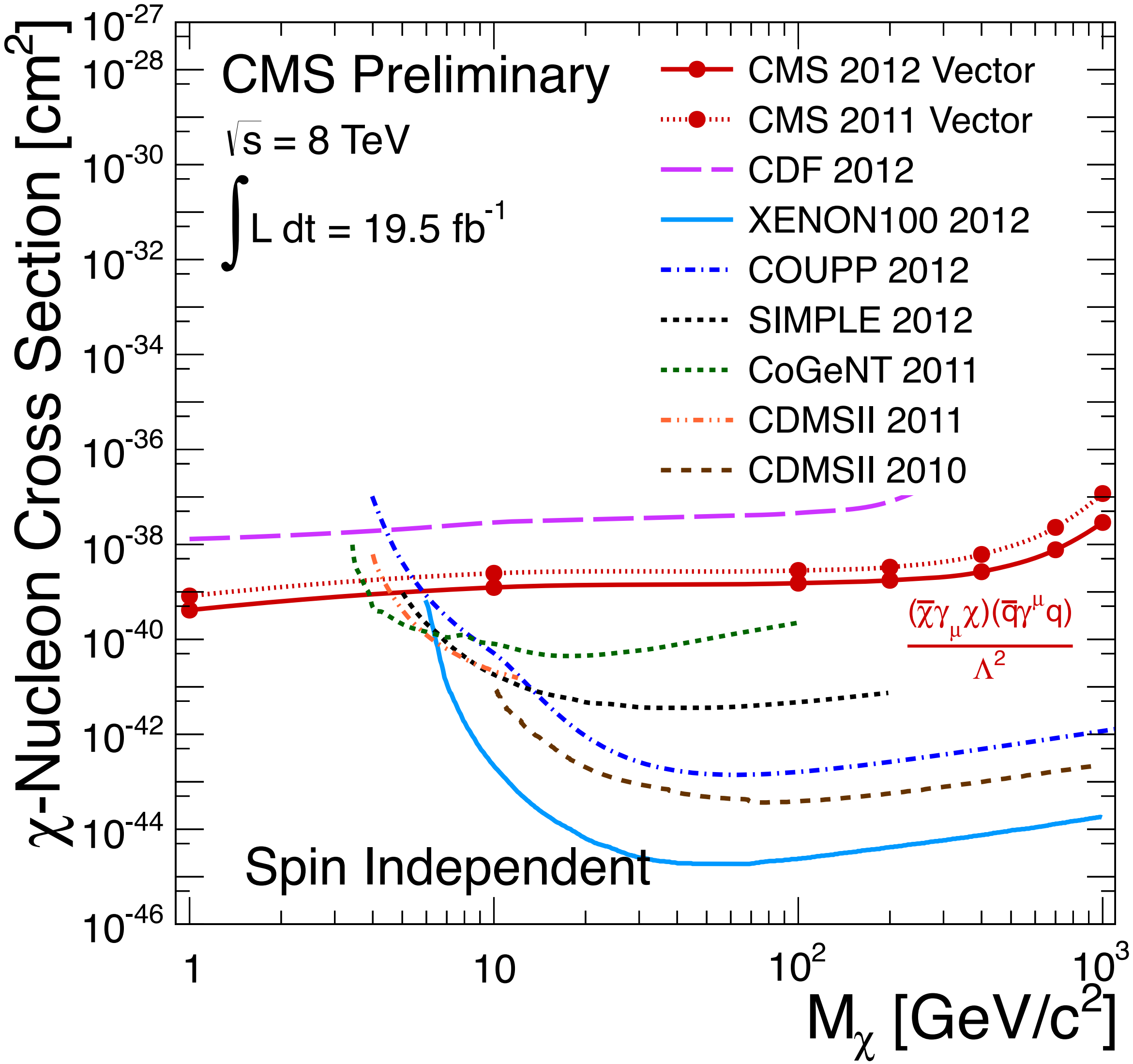


Swamped by QCD. Apply some W/Z tagging requirements on one/ both jets, then set limits on BSM using dijet invariant mass.

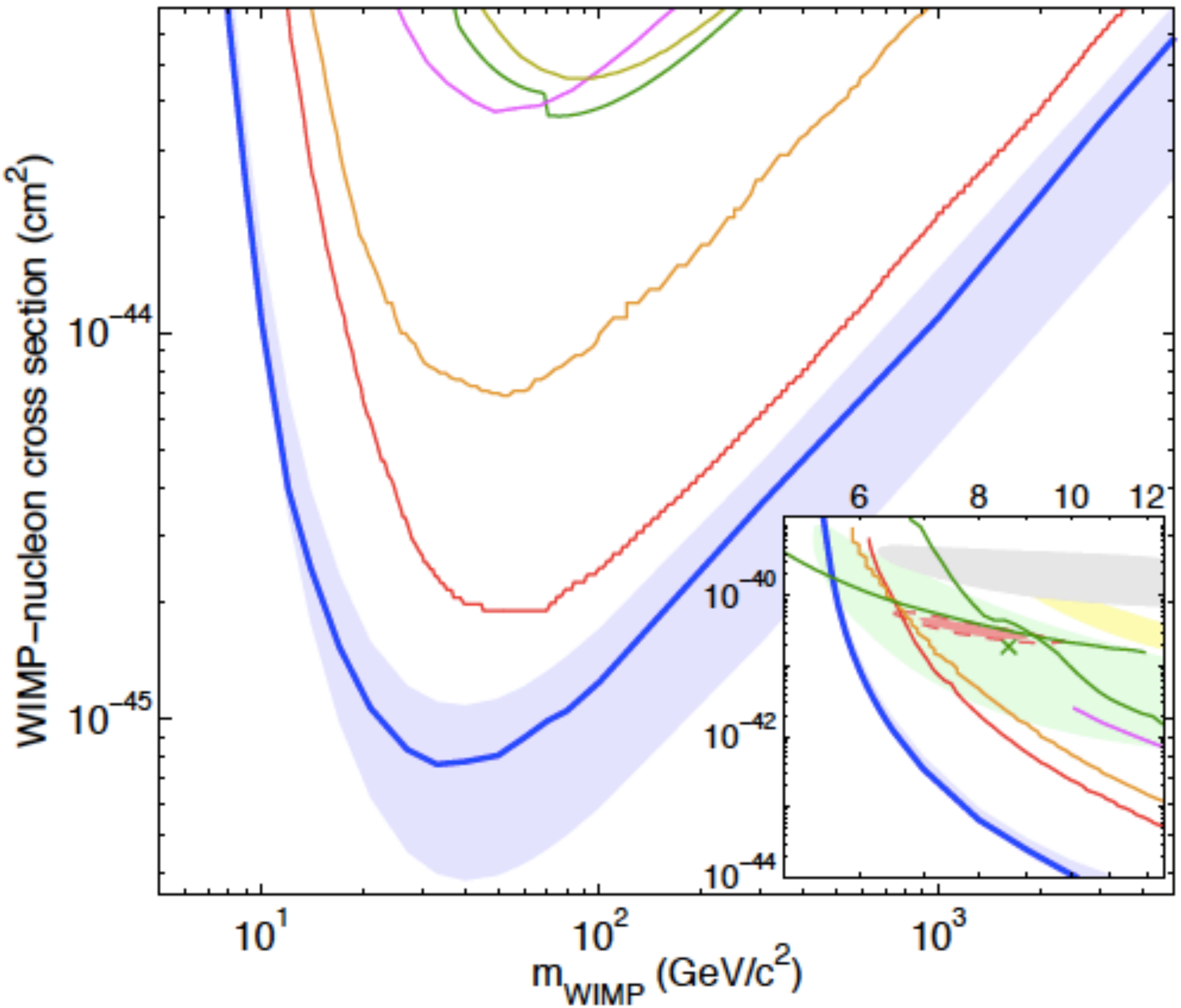
Best limits on low mass WIMP DM particle



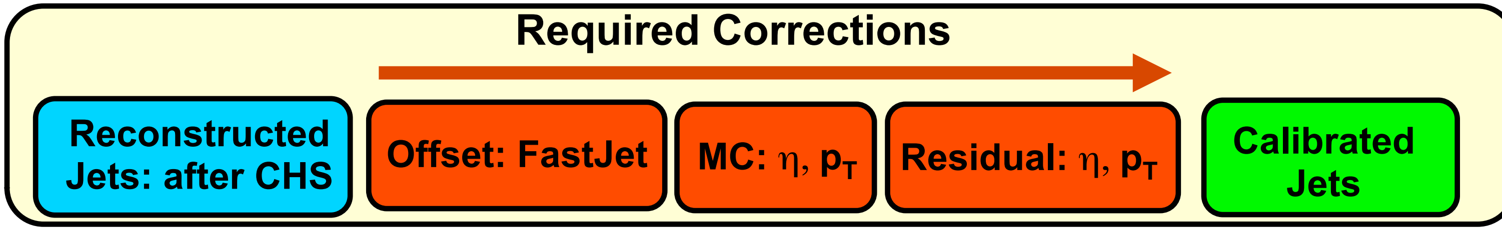
Best limits on low mass WIMP DM particle



Latest result from LUX



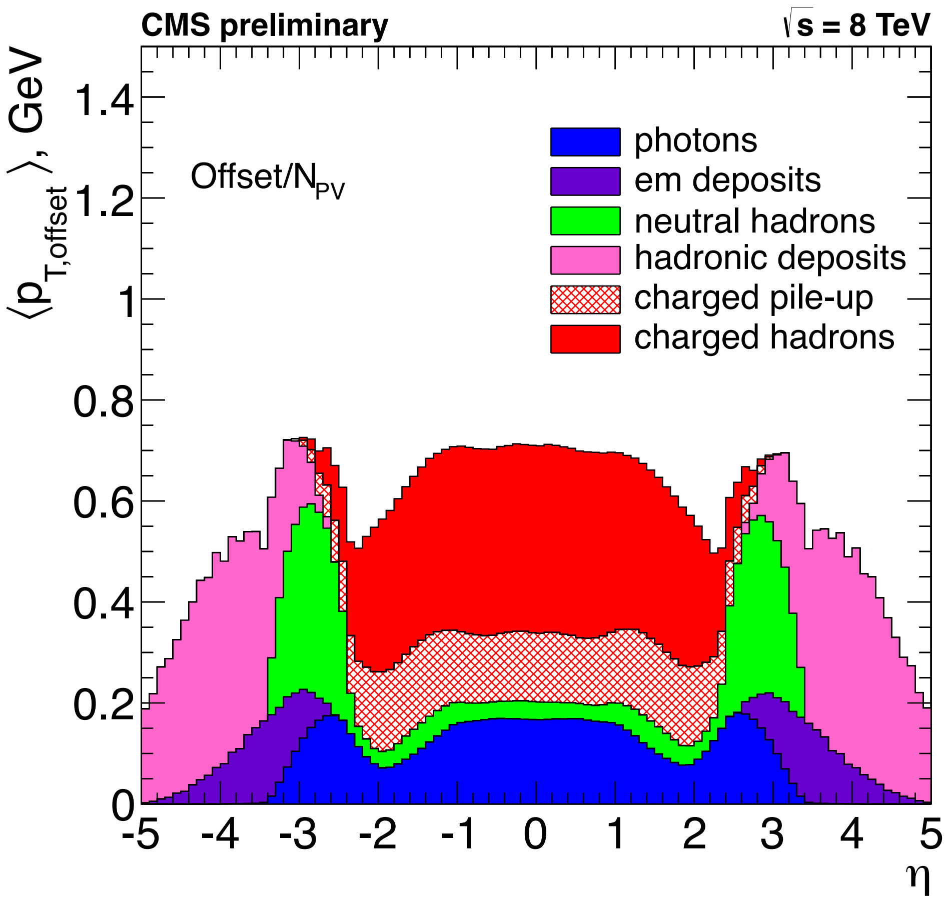
Jet energy calibration: overview



- Factorization facilitates the use of data-driven corrections
  - Breaking the correction into pieces that are naturally measured in collider data:
    - **Offset**: pile-up and noise measured in zero-bias events.
    - **MC**: jet response vs. η, P_T using MC truth.
    - **Residual**: jet response vs. η, P_T using dijet balance and γ/Z+jet in data.

In CMS the most widely used jet is anti-k_T 0.5 (0.7 for QCD measurements). Jet substructure studies done with anti-k_T 0.5, 0.7, 0.8 with various grooming techniques.

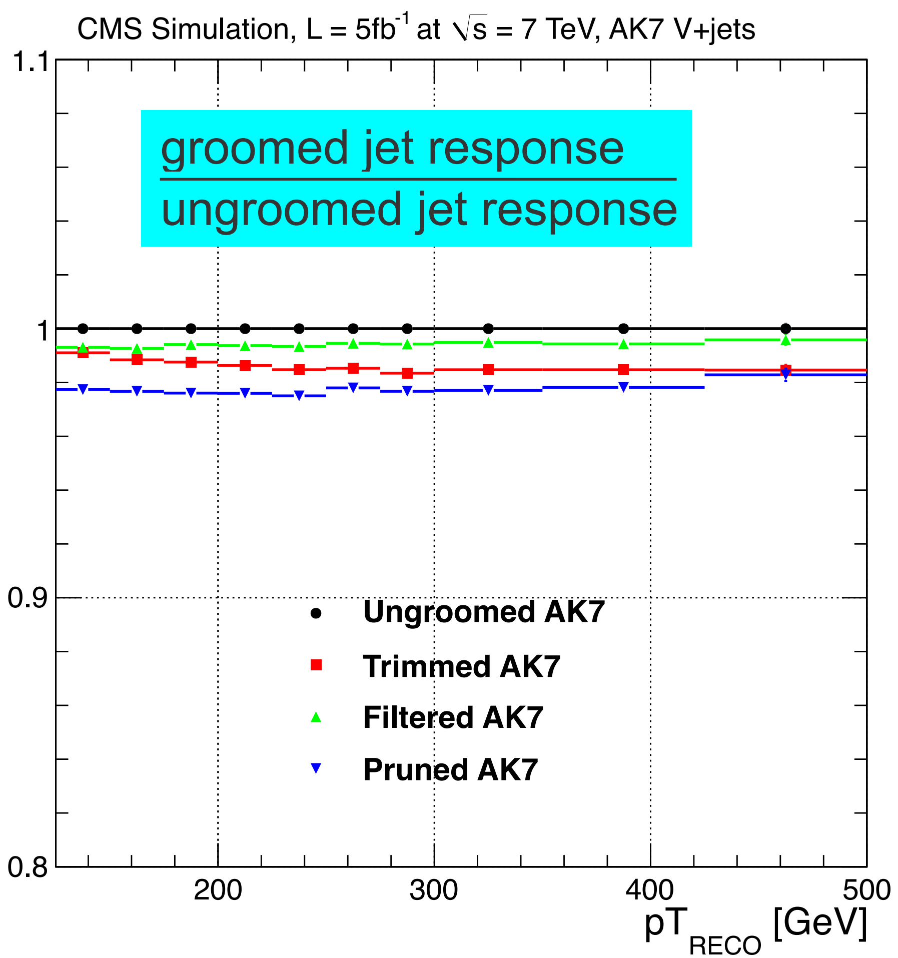
Pileup contribution to jet energy

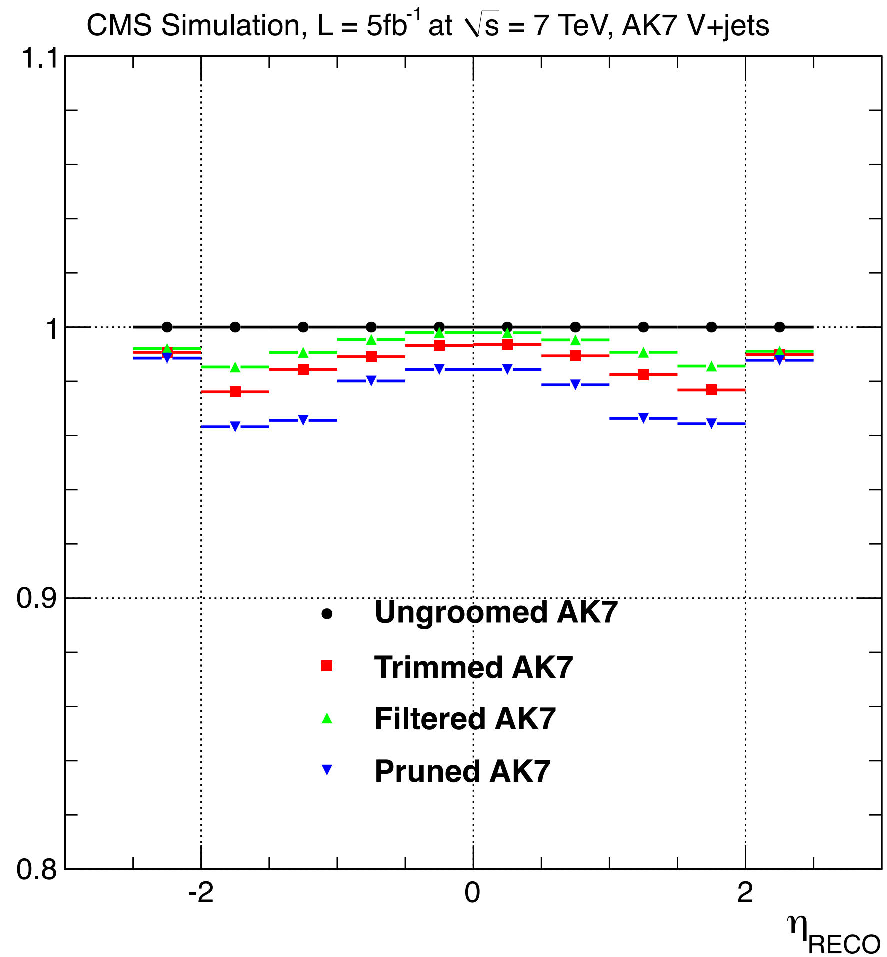


- ✦Pileup (PU) measured with Zero Bias data
  - Most charged hadrons can be associated to pileup vertices and removed
  - Part that can be removed is labeled "charged hadrons"
  - Part that remains as PU needs to be subtracted
    - PU density x effective area (FastJet-ρ)

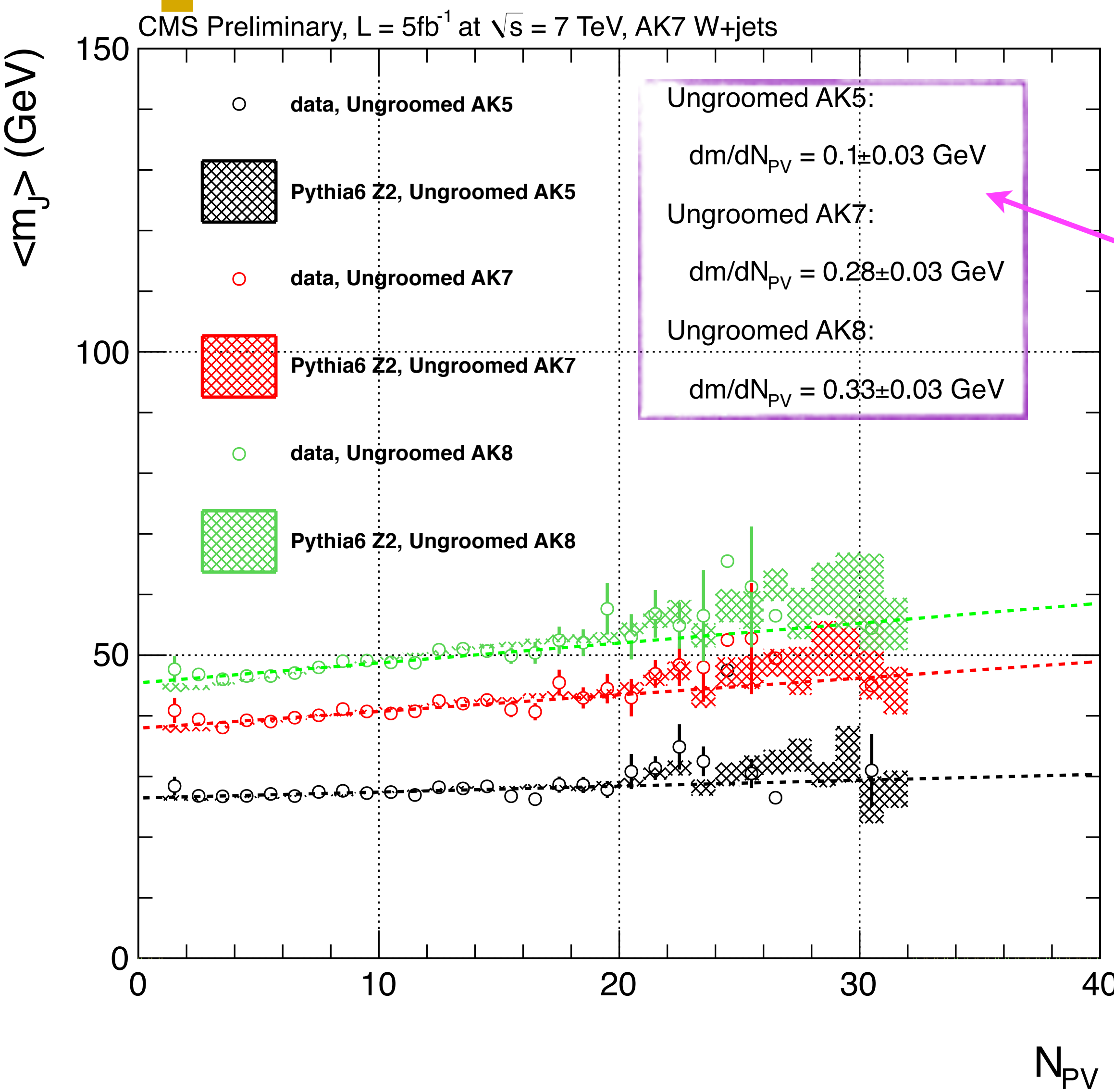
Jet p_T response for groomed jets

Groomed jet response within a few % of ungroomed case.





Performance versus pileup by jet size



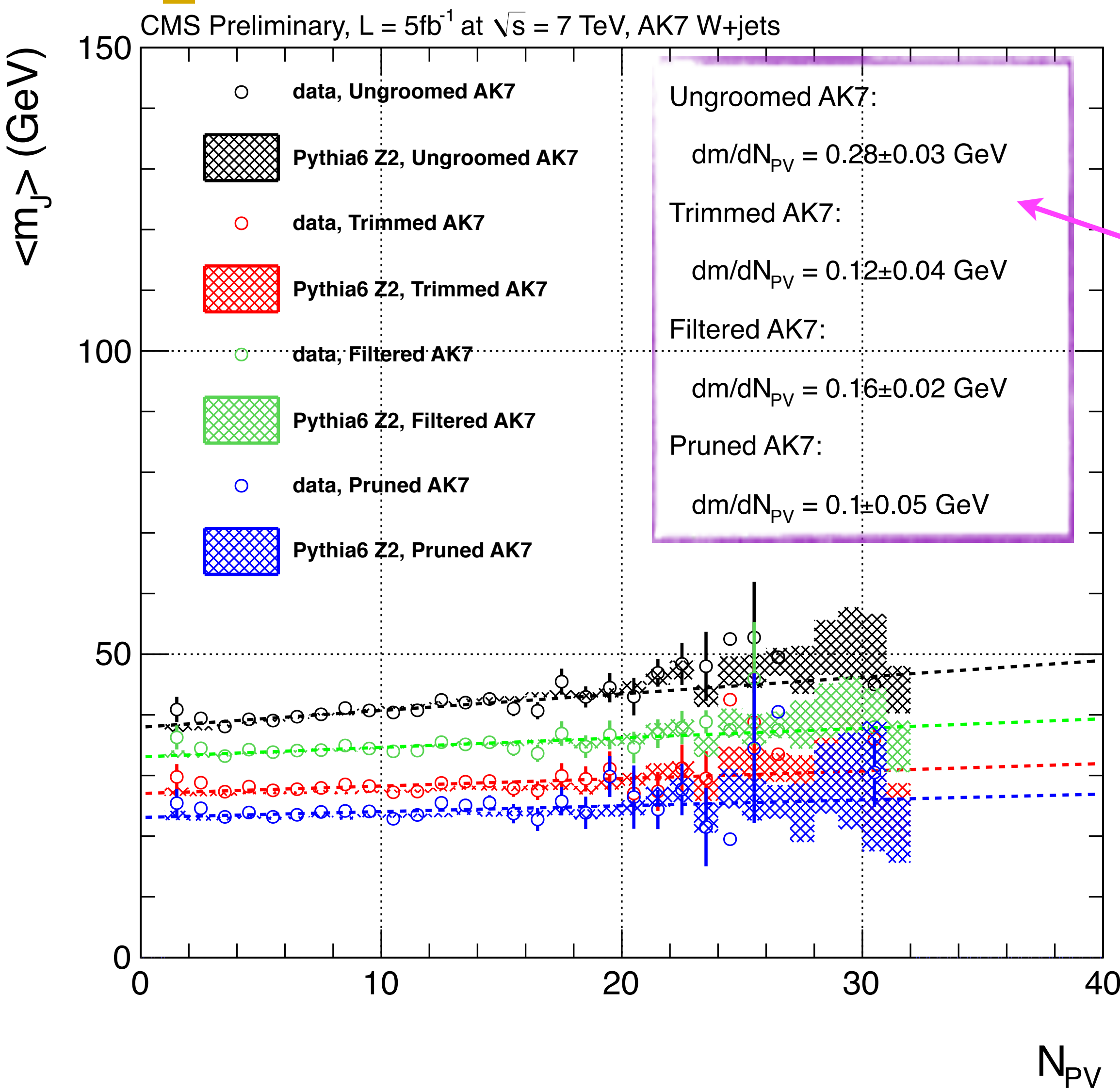
✦Ungroomed jet mass is very sensitive to PU

- <mJ> increases linearly as a function of the number of primary vertices

✦Effect becomes more pronounced as the jet size increases

- AK8 shows much worse effect than AK5

Performance versus pileup for groomed jets



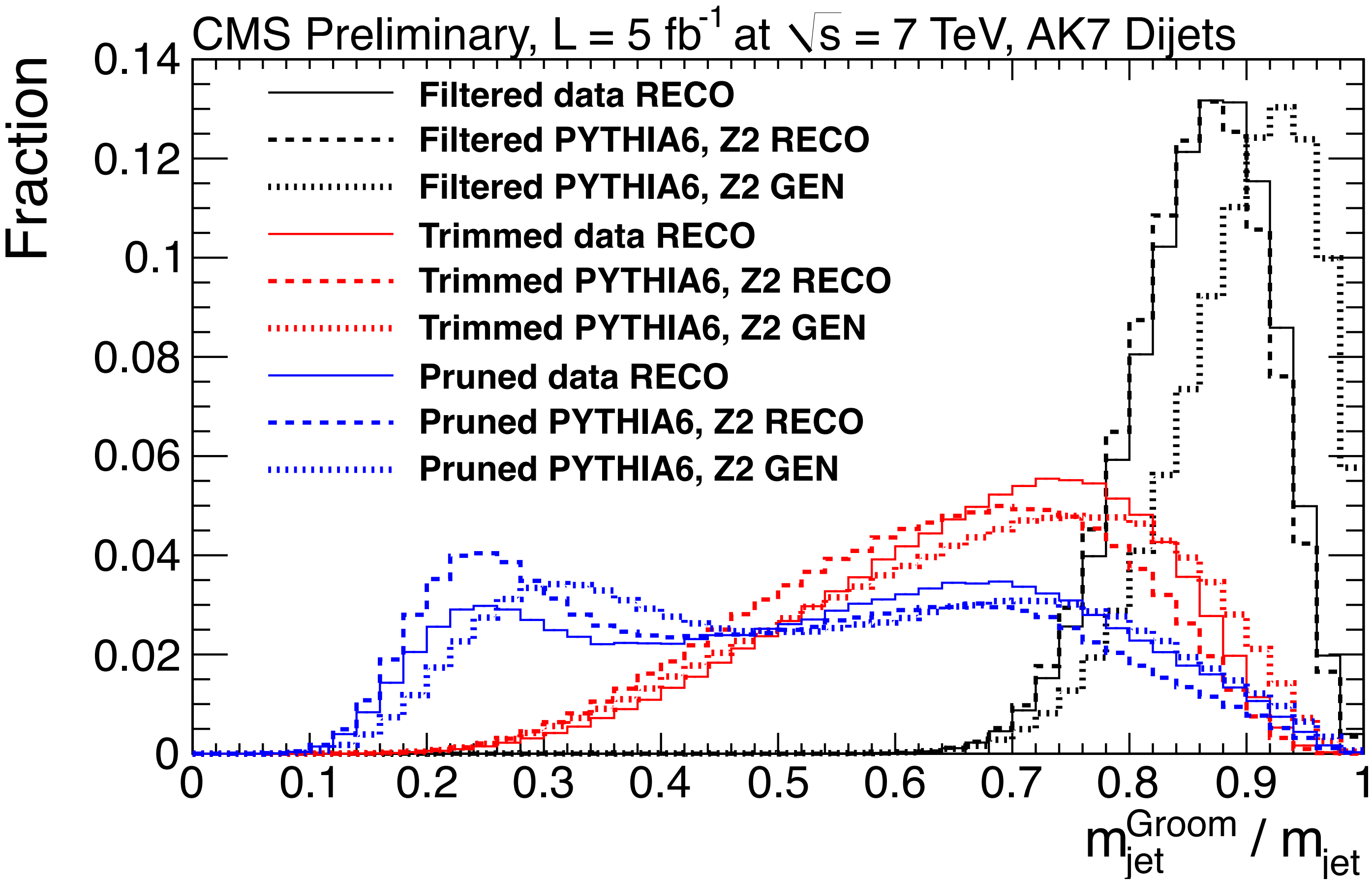
- ✦Grooming techniques are less sensitive to PU
  - •$\langle m_J \rangle$ vs NPV slope becomes flatter
- ✦Observe the expected behavior that $\langle m_J \rangle$ typically scales as R^3

$s_{0.7}/s_{0.5}$	=	2.7 ± 0.9	$((0.7/0.5)^3 = 2.74)$,
$s_{0.8}/s_{0.5}$	=	3.3 ± 1.0	$((0.8/0.5)^3 = 4.10)$,
$s_{0.8}/s_{0.7}$	=	1.2 ± 0.2	$((0.8/0.7)^3 = 1.49)$

Jet mass for groomed jets in background

arXiv:1303.4811

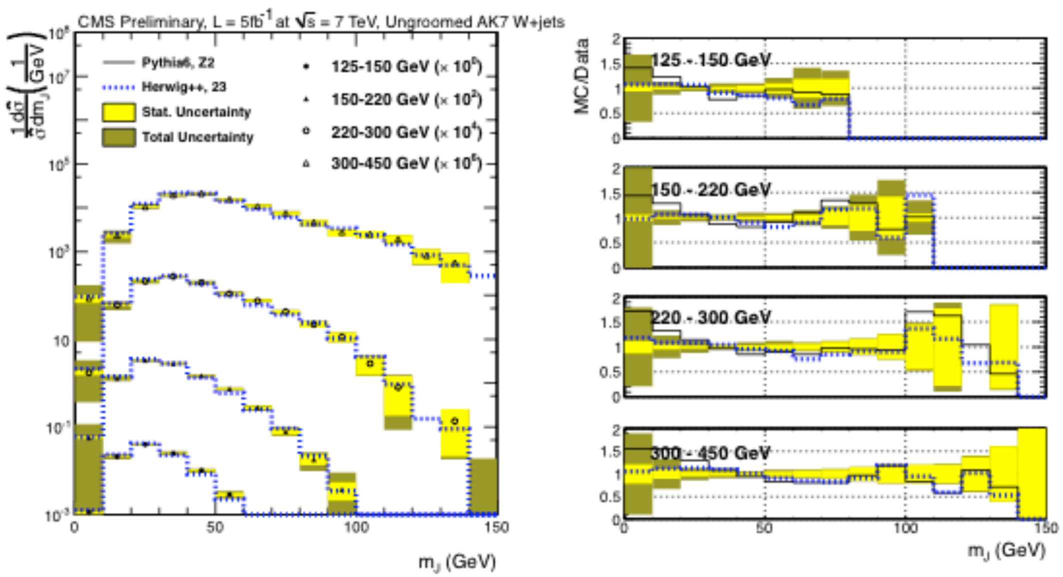
Ratio of the groomed to ungroomed jet mass



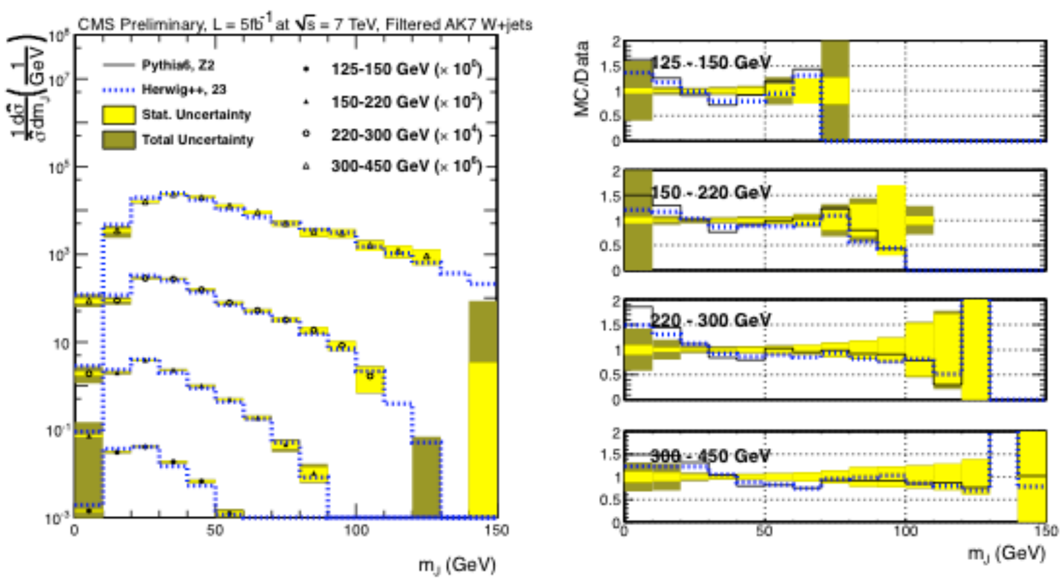
Comparison of grooming algorithms at particle level (GEN), reconstructed simulation (RECO) level, and in data

Unfolded distributions, W+jet (AK7)

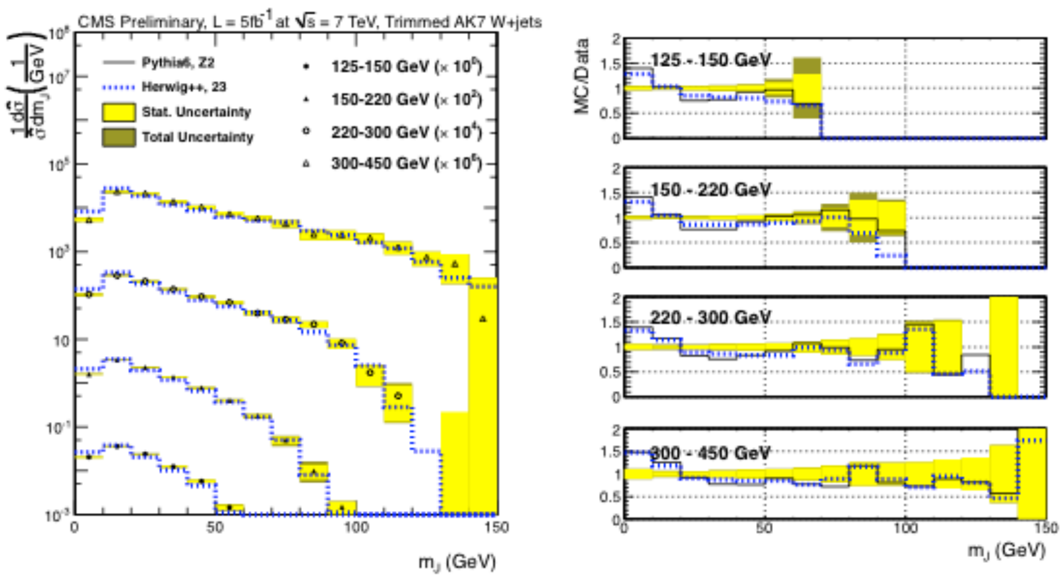
AK7, UNGR



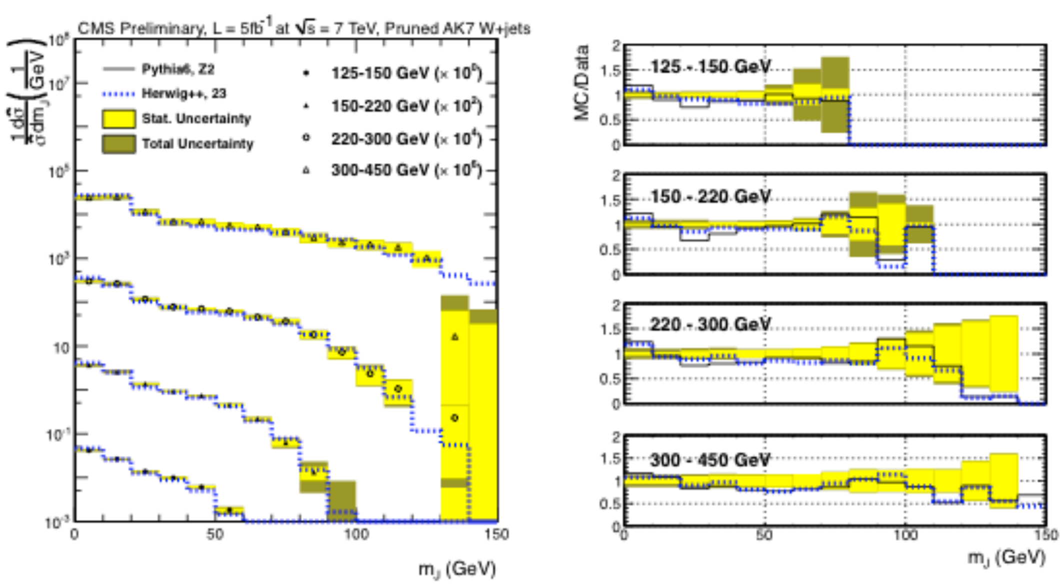
AK7, Filtered



Unfolded distributions, W+jet (AK7)

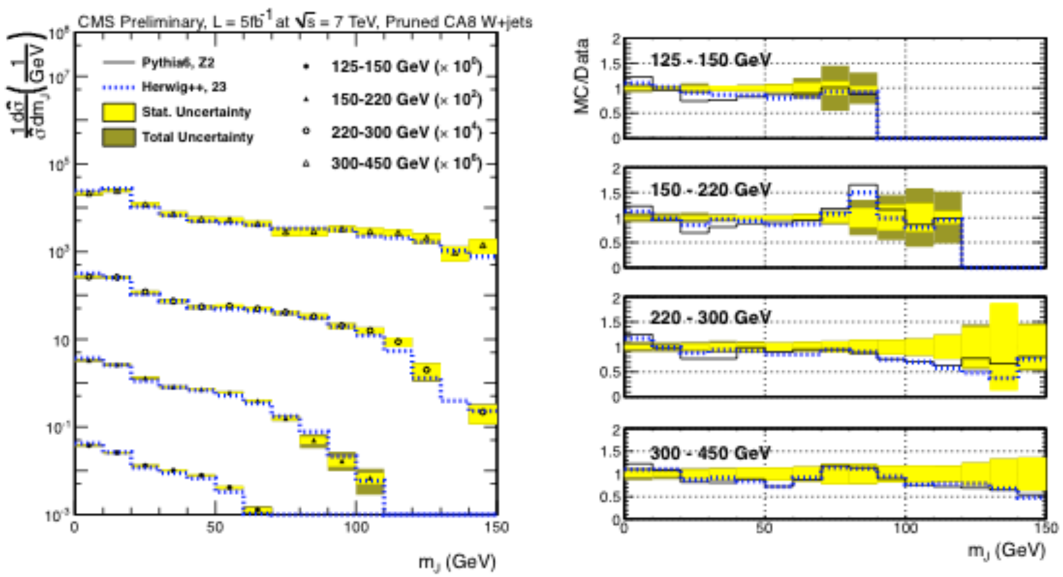


AK7, Trimmed

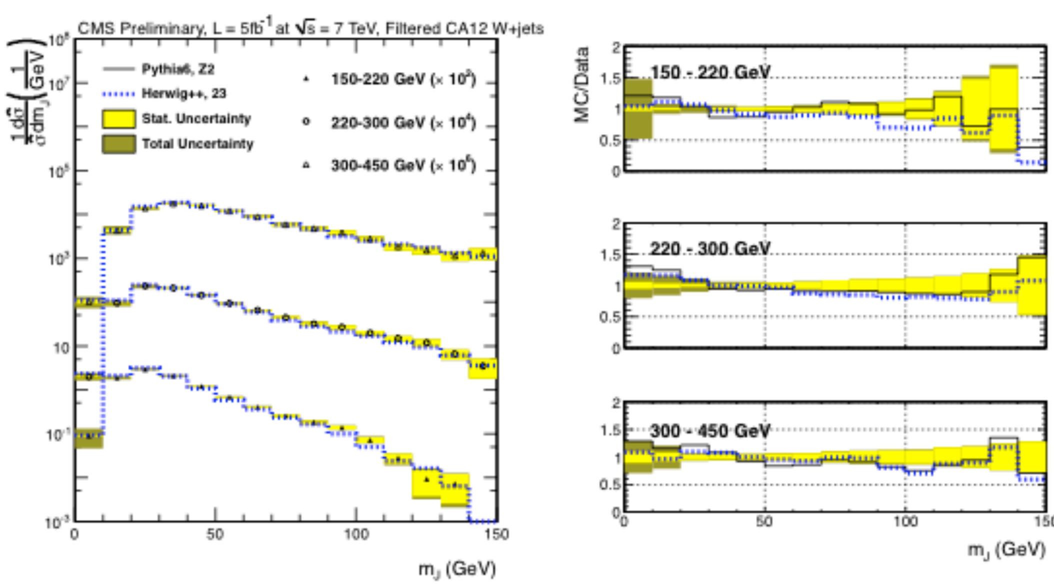


AK7, Pruned

Unfolded distributions, W+jet: comparison



CA8, Pruned



CA12, Filtered