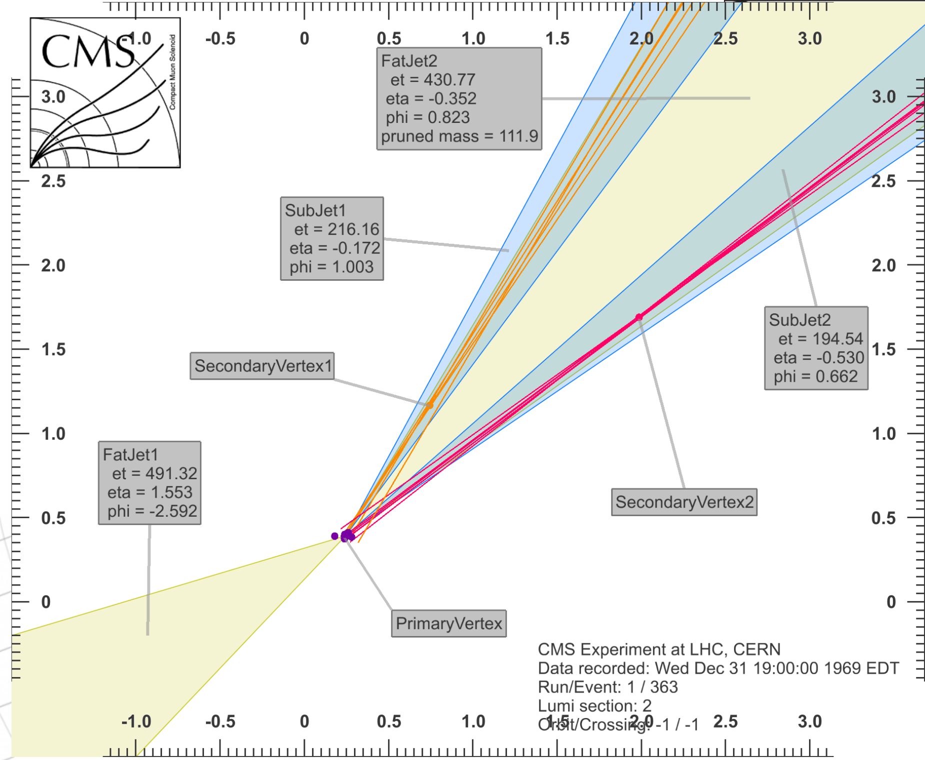
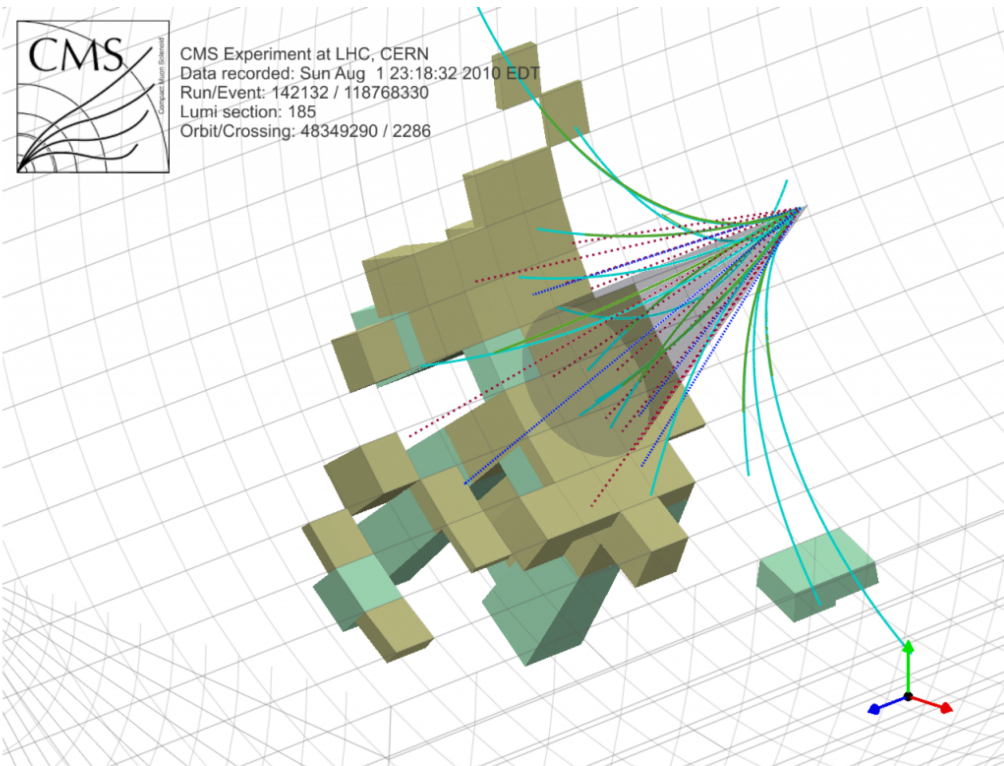
BSM searches in CMS using 'boosted' jets





Devdatta Majumder

University of Kansas

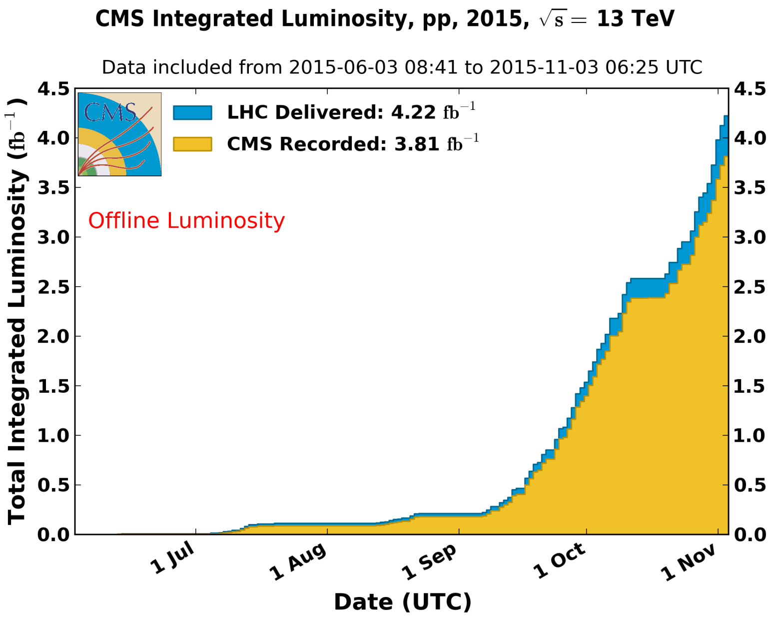
and the CMS Collaboration

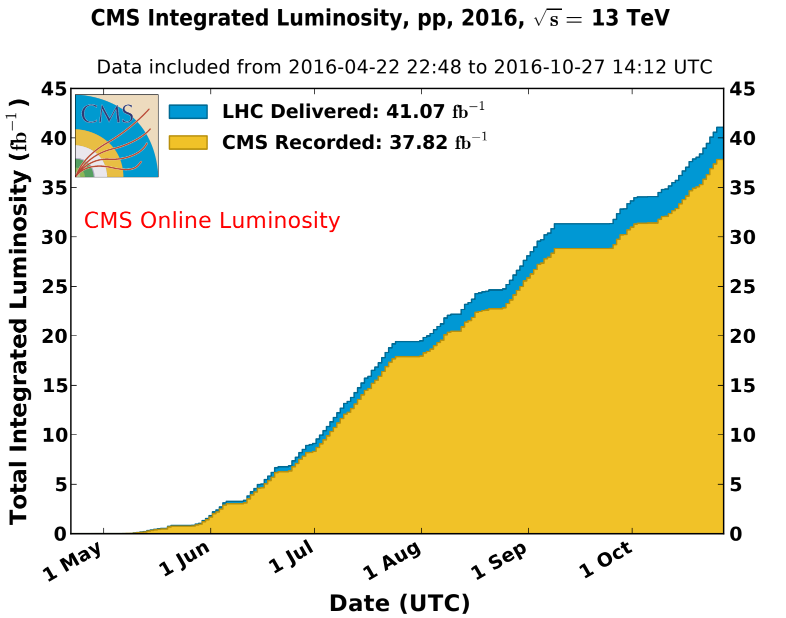
Jets@LHC Workshop

ICTS, TIFR

Bangalore, 27 Jan 2017

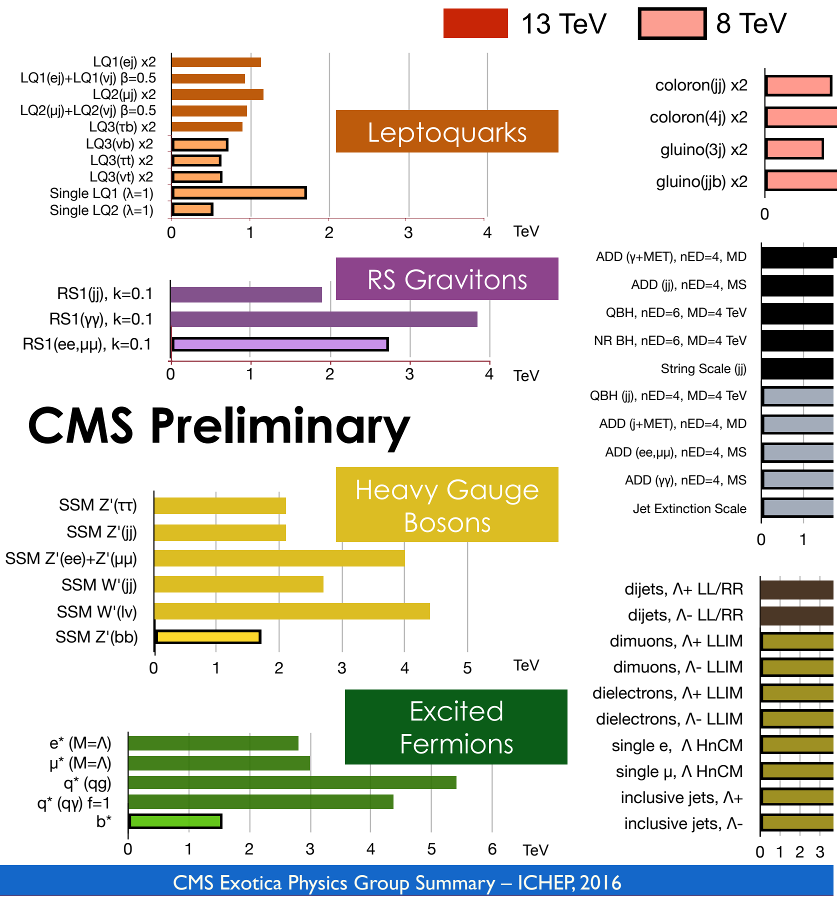
LHC performance



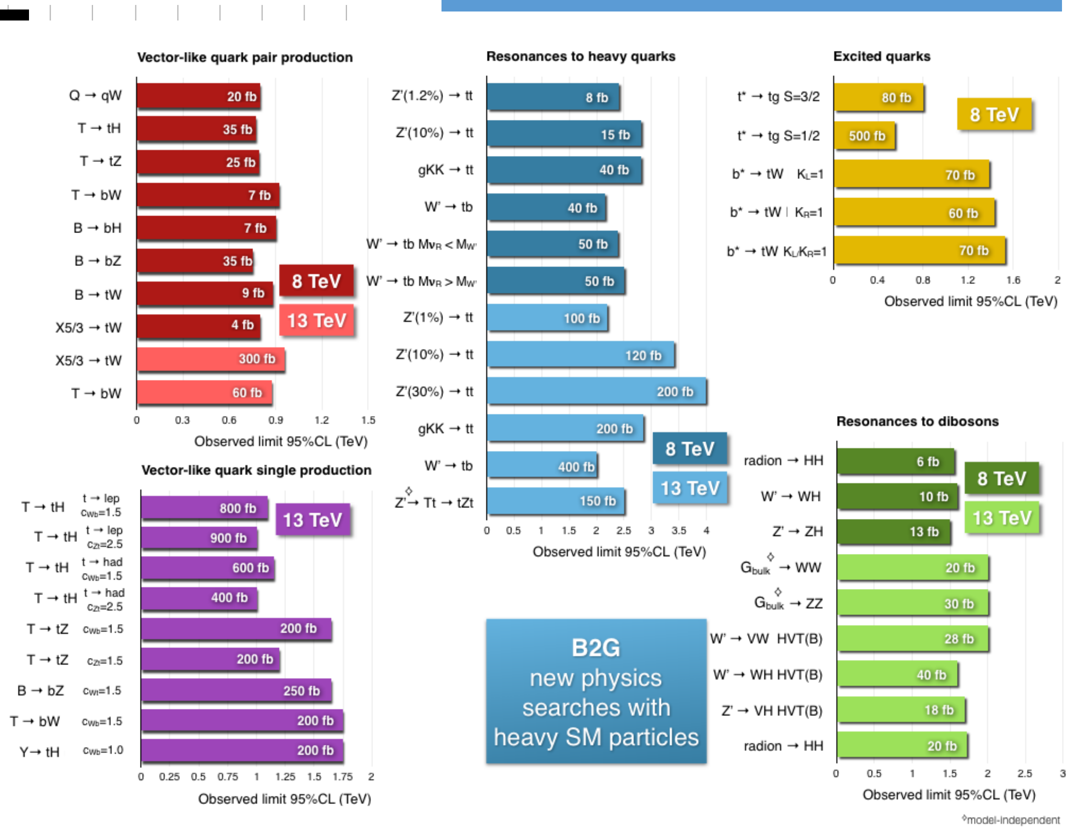


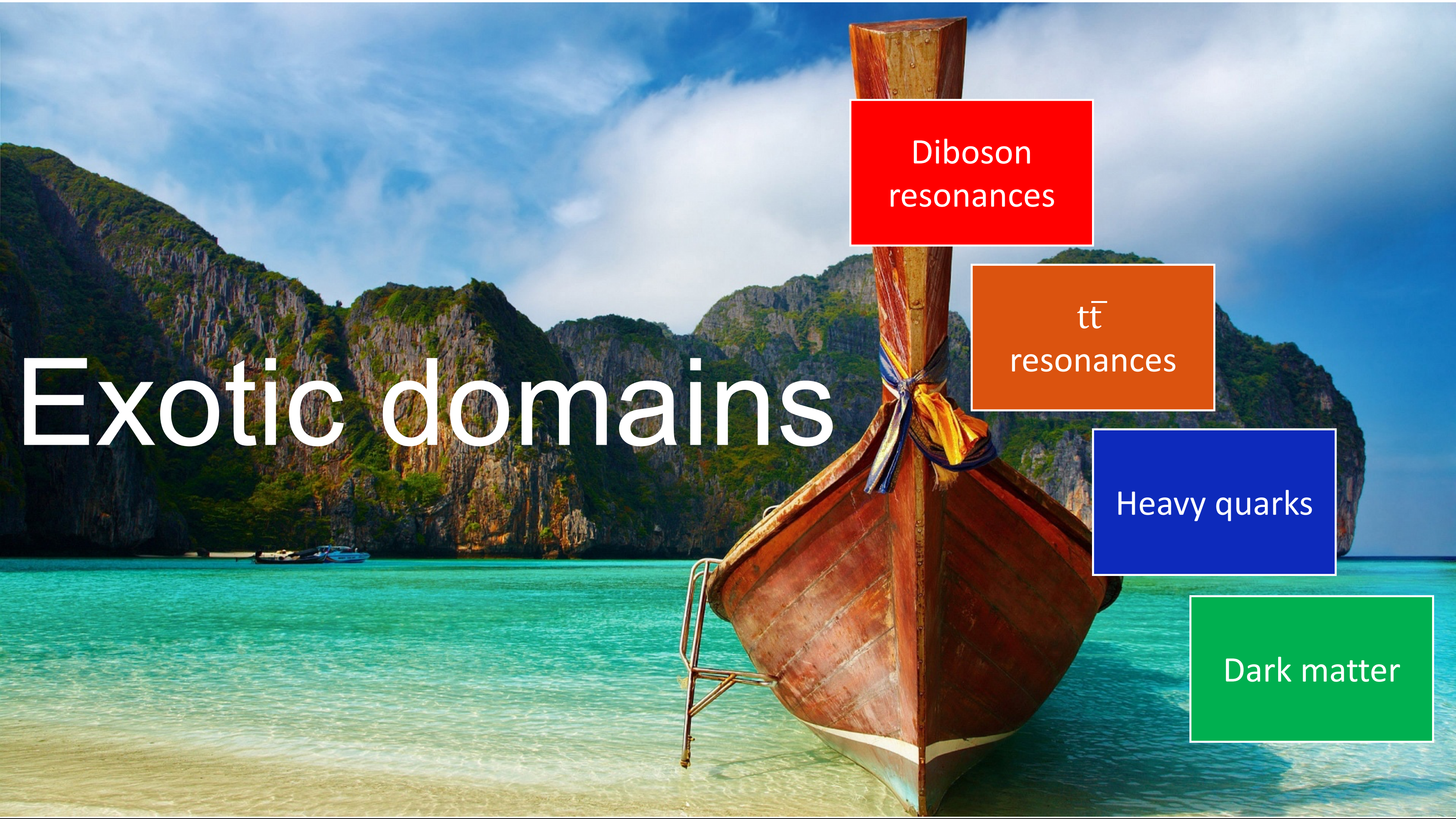
- Run 2 has been amazing so far!

CMS searches in a nutshell



Searches being pushed to **higher masses**



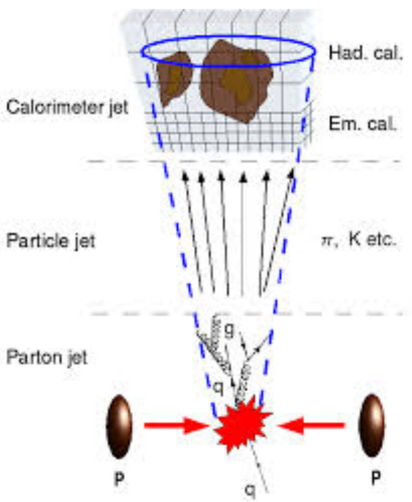


Exotic domains

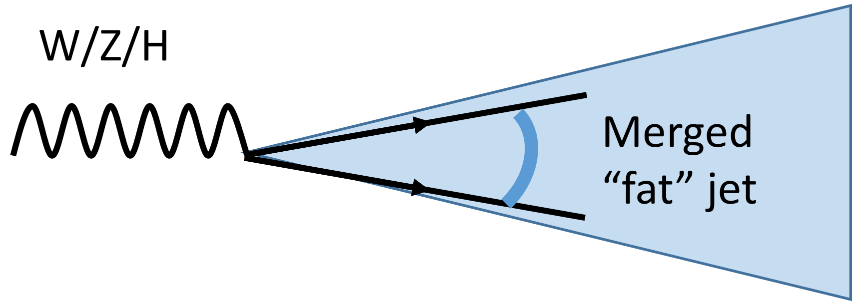
- Diboson resonances
- $t\bar{t}$ resonances
- Heavy quarks
- Dark matter

The toolbox

Jets



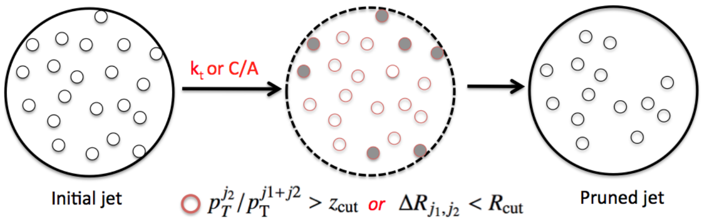
W/Z/H



$$\Delta R \approx \frac{2M}{p_T}$$

❑ Pruning:

❖ $R_{cut} = 0.5\frac{m}{p_T}$ $z_{cut} = 0.1$



❑ Soft drop (modified mass drop): $\frac{\min(p_{T1},p_{T2})}{p_{T1}+p_{T2}} > z_{cut}\left(\frac{\Delta R_{12}}{R}\right)^{\beta}$

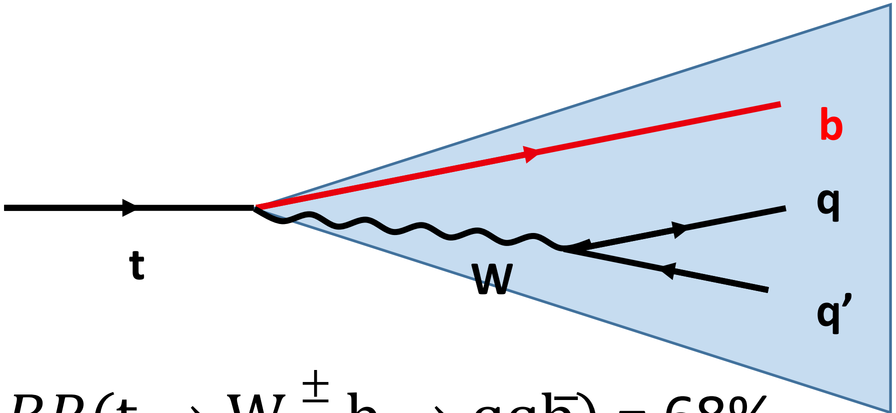
❖ $\beta = 0$ $z_{cut} = 0.1$

❑ N-subjettiness $\tau_N = \frac{1}{d_o}\sum_k p_{T,k}\min\{\Delta R_{1,k}, \Delta R_{2,k}, \ldots, \Delta R_{N,k}\}$

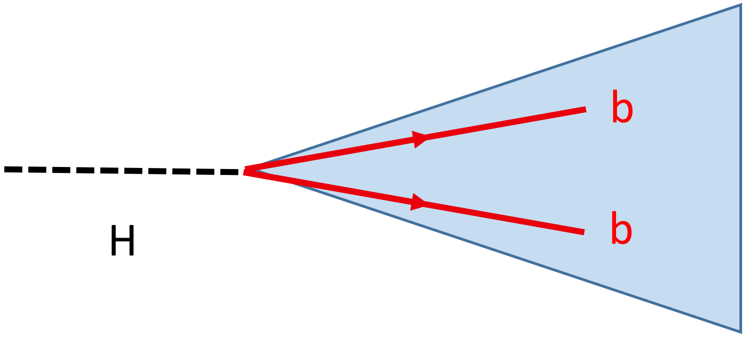
❖ τ_2/τ_1 for W/Z/ Higgs tagging

❖ τ_3/τ_2 for top tagging

B-tagging in boosted topologies



$BR(t \to W^{\pm} b \to qq\bar{b}) = 68\%$

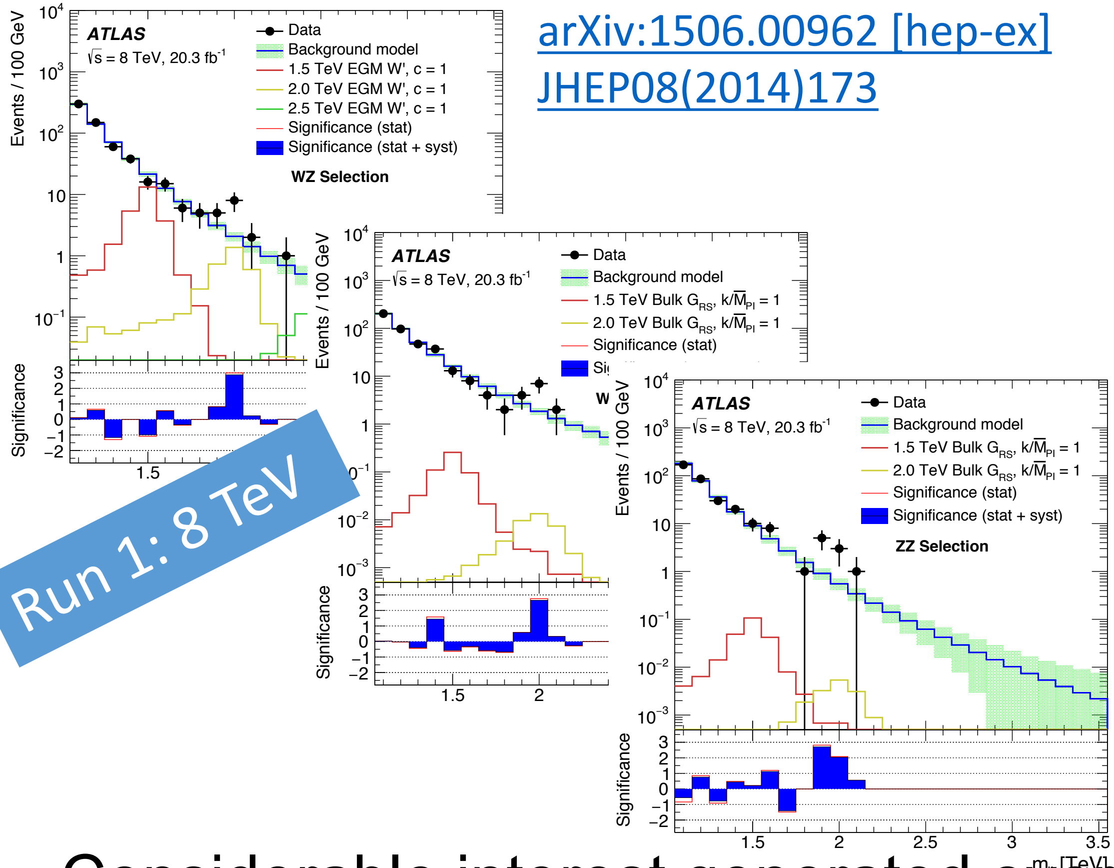


$BR(H \to b\bar{b}) = 58\%$

- Boosted top quark and Higgs bosons form final state of many BSM particle decay.
  - Use fully hadronic decay to enhance signal
  - B-tagging subjets improve signal sensitivity

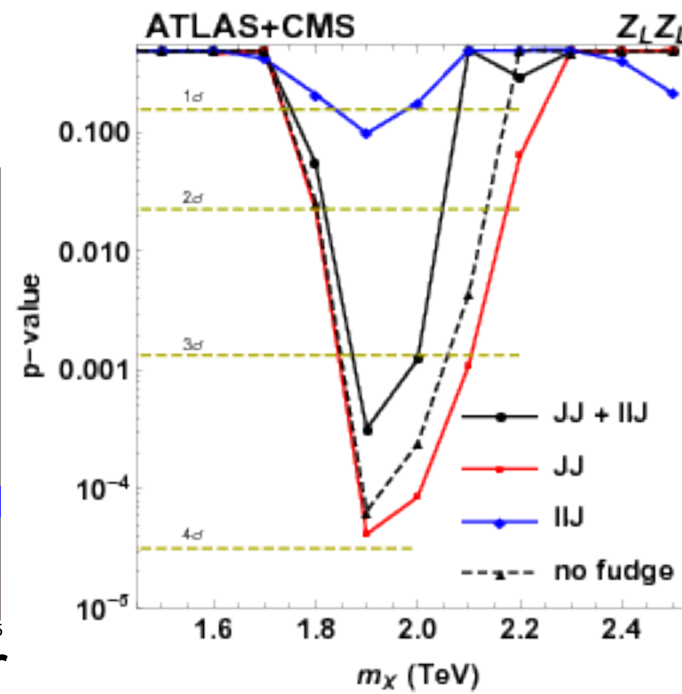
Diboson resonances

The Run 1 excess



arXiv:1506.00962 [hep-ex]
JHEP08(2014)173

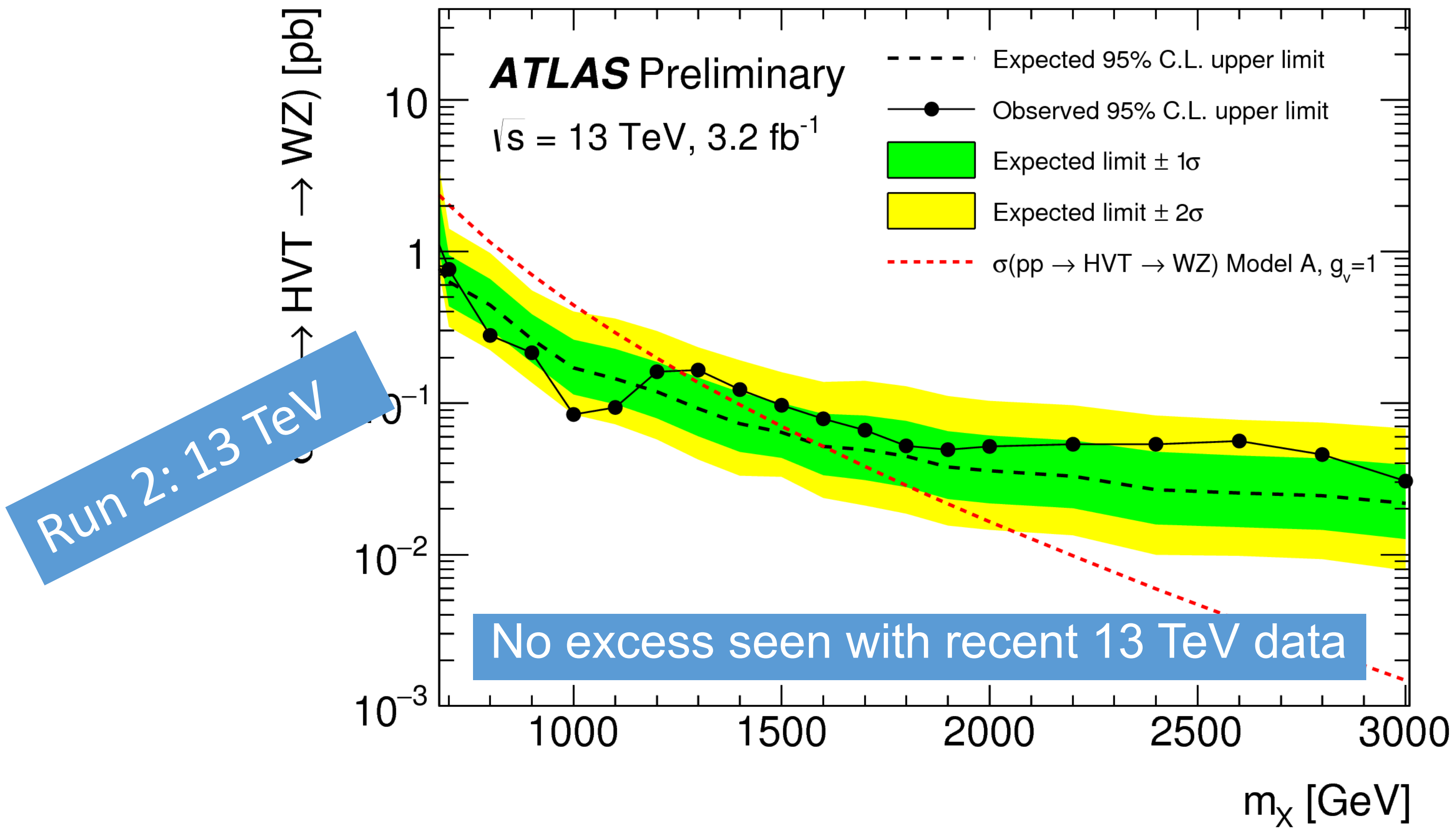
ATLAS+CMS combination shows pretty big "evidence" for an excess



JHEP04(2016)155

- Considerable interest generated over time.

And then Run 2 started….

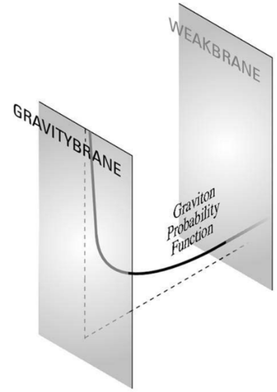


M. Kado/CERN/15Dec2015

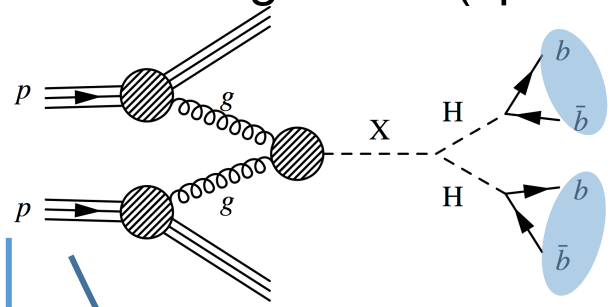
Extradimension models

- Warped extradimension models as solution to the hierarchy problem.
- SM fields can propagate in the "bulk".

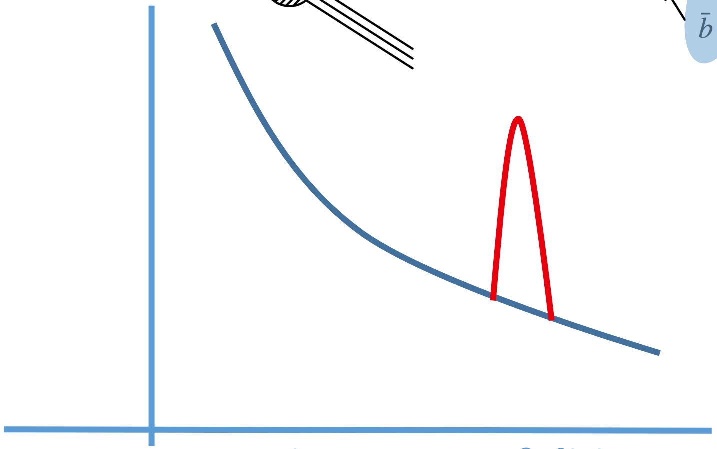
Plank scale $\overline{M_{\mathrm{Pl}}}$
Warp factor κ
Mass scale
$\Lambda_R = \sqrt{6}e^{-\kappa l}\times\overline{M_{\mathrm{Pl}}}$
~ TeV (κl ~ 35)



- Additional degrees of freedom
  - Radion (spin 0)
    bulk graviton (spin 2)



- Search strategy: Bump hunt over a smooth background
  - Mostly QCD dijet production

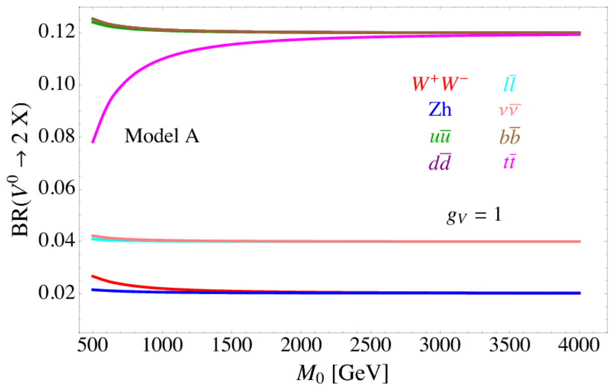


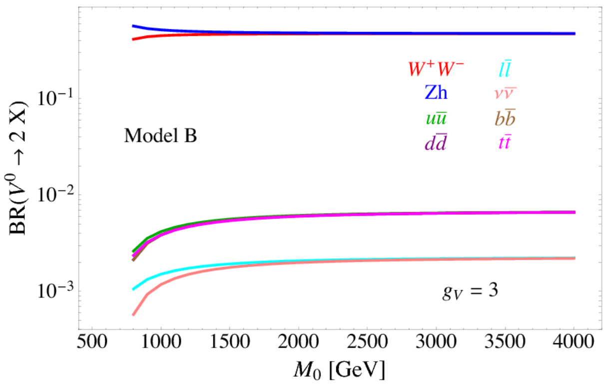
The heavy vector triplets

- Extension of the SM gauge sector
- W' and Z': Heavier partners of the SM W and Z bosons
- Comparable coupling to quarks and SM bosons – Model A
- Mostly coupling to SM bosons – Model B
  - Accessible through diboson resonance searches
  - Resonance mass, couplings to fermion g_F and gauge bosons g_V a

For no quark couplings, accessible only through VBF productions

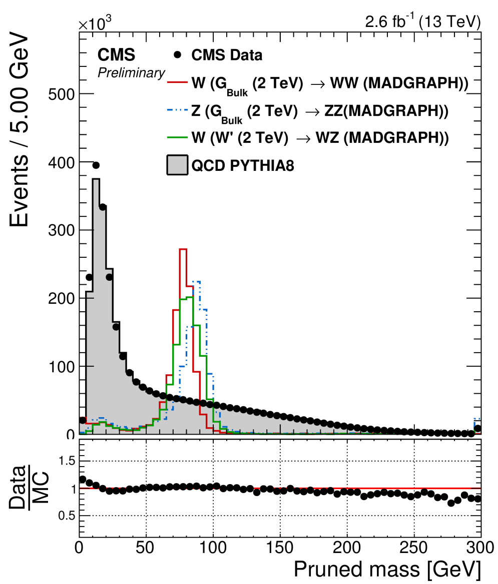
JHEP 1409(2014) 060

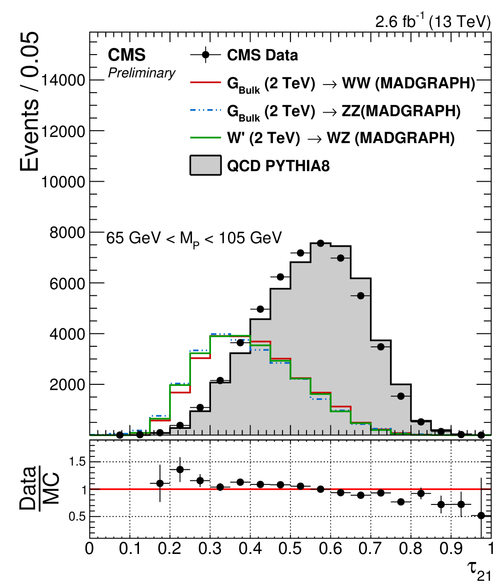




Picking the signal

- Narrow resonances – peaking signal
  - Jet substructure to reduce multijets backgrounds
  - Jet groomed mass and N-subjettiness main workhorses
  - $\tau_2/\tau_1 < 0.45$ or $0.45 < \tau_2/\tau_1 < 0.75$
  - 65 < pruned mass < 105 GeV



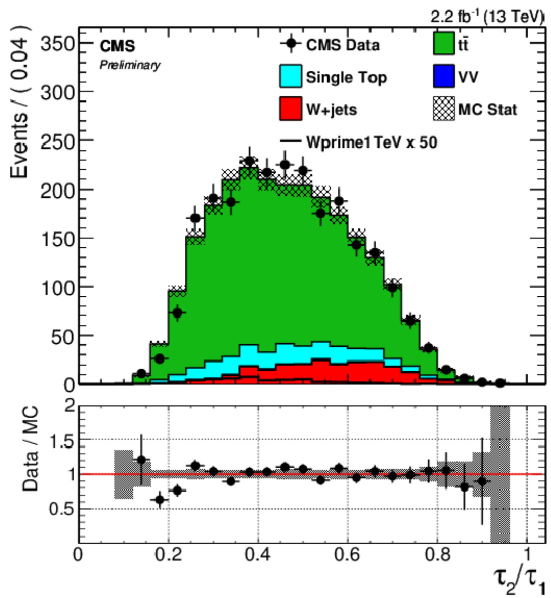


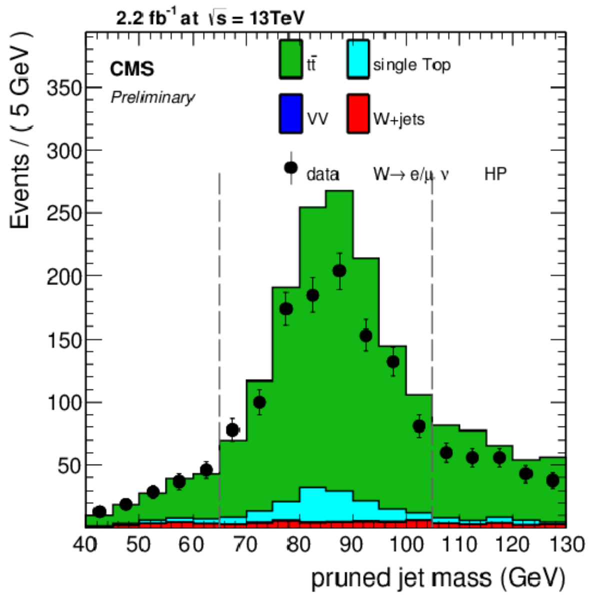
CMS-EXO-15-002

Measurement in the data

- W boson jets from $t\bar{t}$ events used for comparing simulation to the data

CMS-EXO-15-002





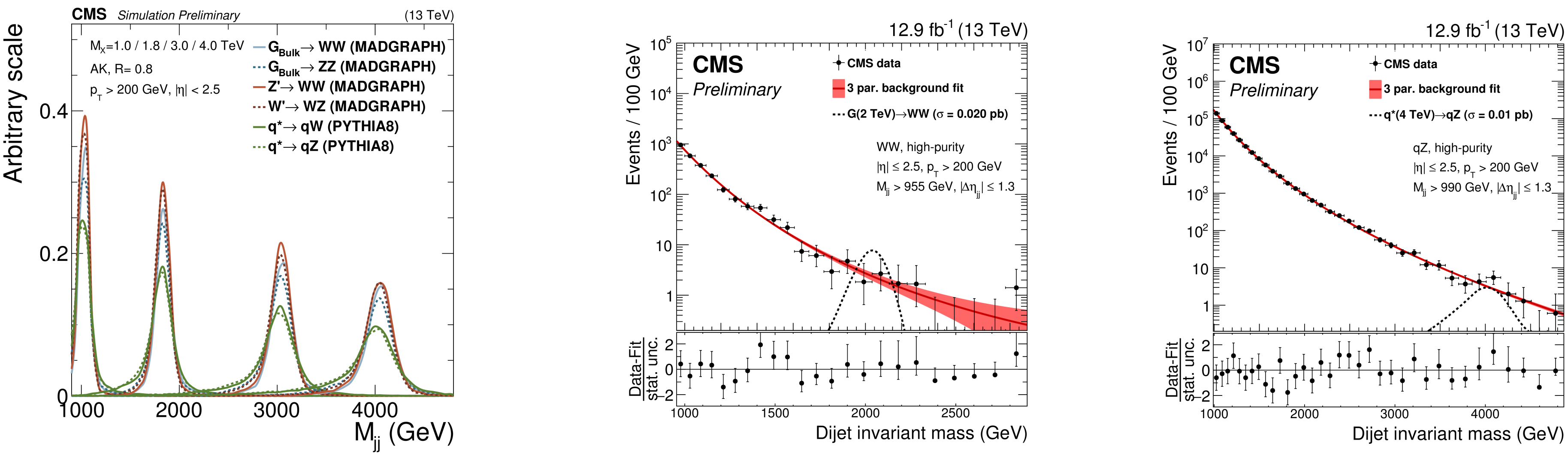
Category	Definition	W scale factor
Dijet channel HP	$(\tau_{21} < 0.45)$	0.69 ± 0.14
Dijet channel LP	$(0.45 < \tau_{21} < 0.75)$	1.46 ± 0.38
$\ell\nu$+jet channel HP	$(\tau_{21} < 0.6)$	1.03 ± 0.13
$\ell\nu$+jet channel LP	$(0.6 < \tau_{21} < 0.75)$	0.88 ± 0.49

$\tau_{21} < 0.45$	m [GeV]	σ [GeV]
Data	84.7 ± 0.4 GeV	8.2 ± 0.5 GeV
Simulation	85.3 ± 0.4 GeV	7.3 ± 0.4 GeV

Modelling the qV/ VV processes

- Resoance particle mass reconstructed from dijet pairs
- Several categories
  - WW/ WZ/ ZZ
  - qW/ qZ
- Smoothness test of background using functional fits

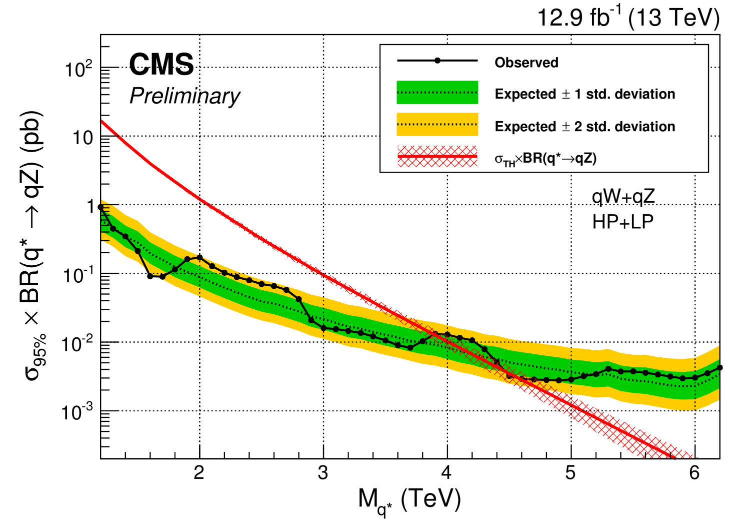
CMS-B2G-16-021

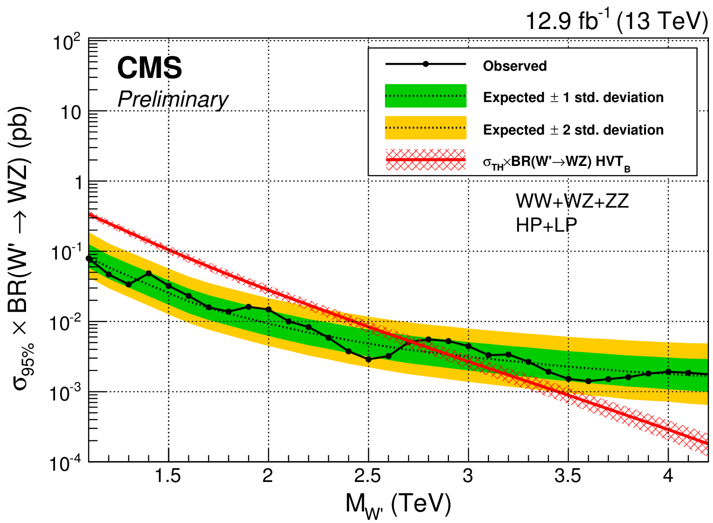


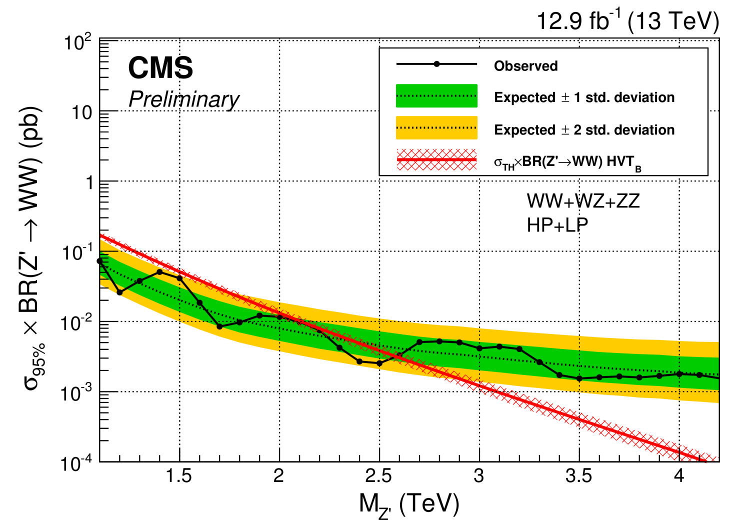
Results on the qV/ VV res. searches

- Limits set on different models in different categories:
  - Bulk gravitons, radions, W', Z', excited quarks

CMS-B2G-16-021



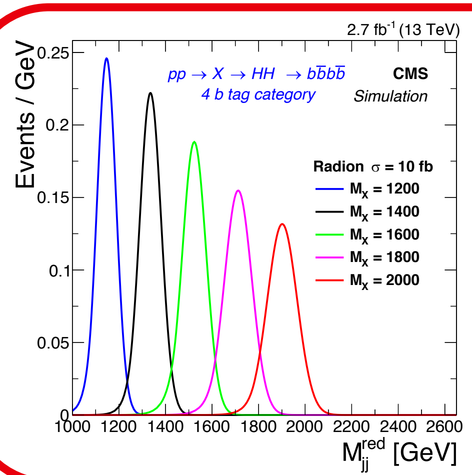


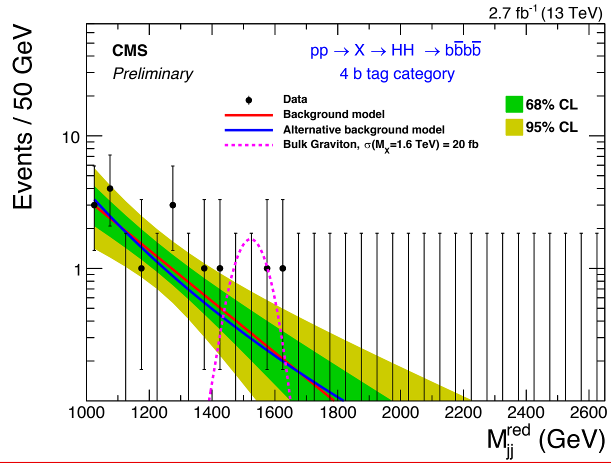


Di-Higgs resonance

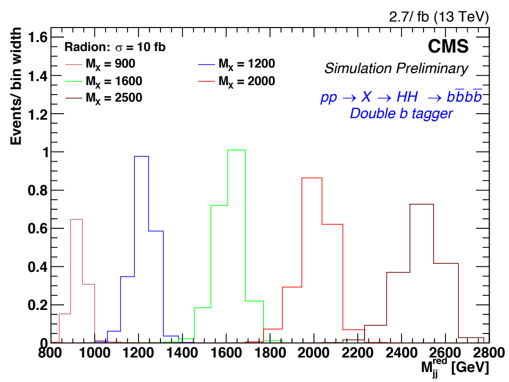
Two analysis approaches:

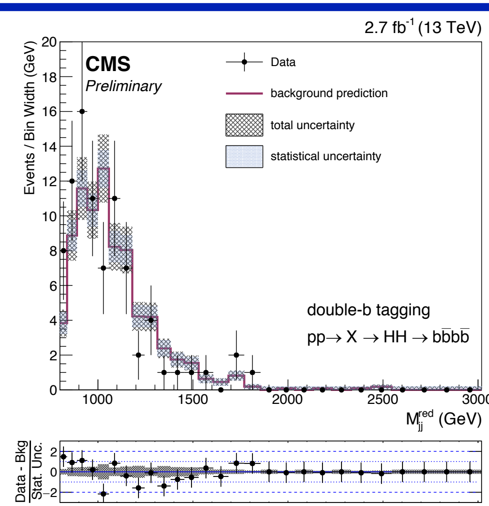
CMS-B2G-16-008





- Parametric signal+background models
- Fit to the data **(smoothness test)**
- Check for excesses

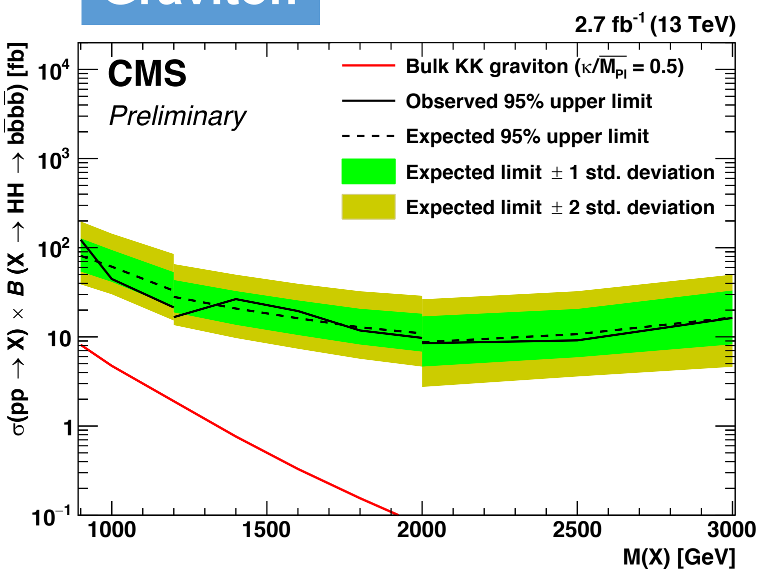




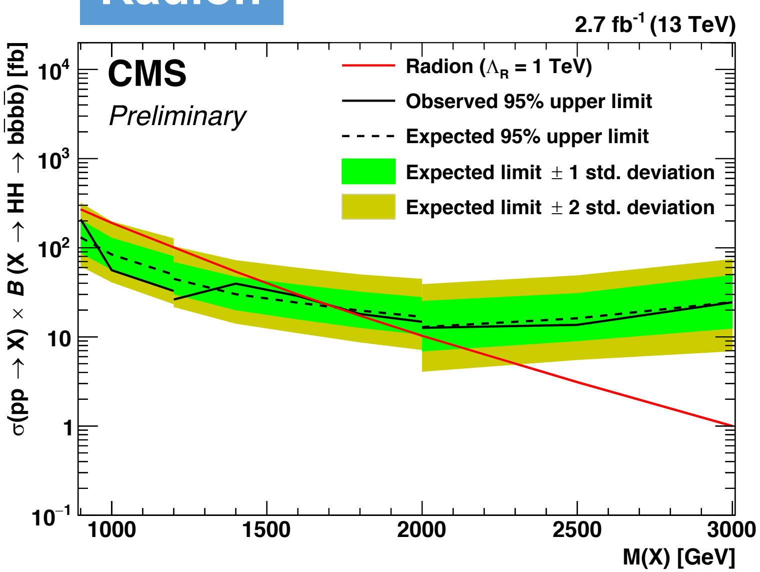
- Exact background estimation based on control regions **("Alphabet" method)**
  - Both normalization and shape

Results on the di-Higgs search

Graviton



Radion

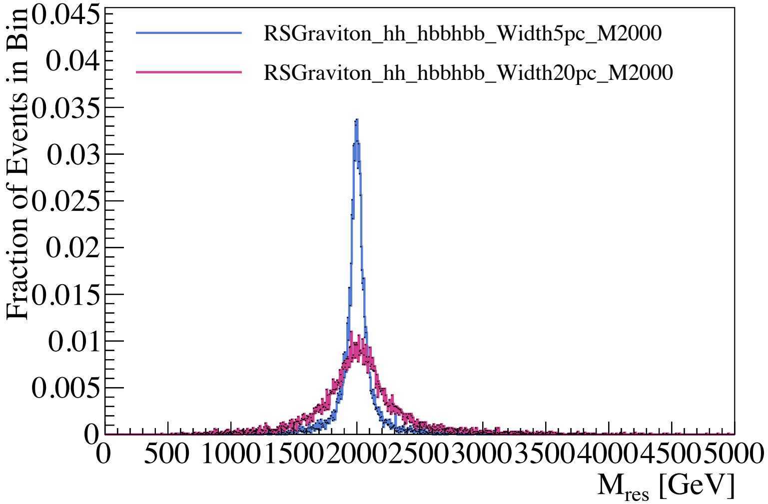


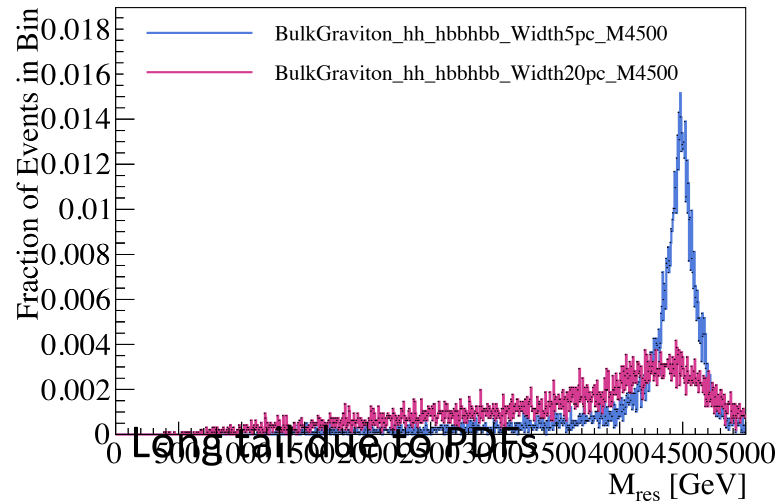
- Low and high mass range – Alphabet method
- Intermediate mass range – Smoothness test

CMS-B2G-16-008

Wide resonances

- Wide resonance search goes beyond traditional “bump hunt”.



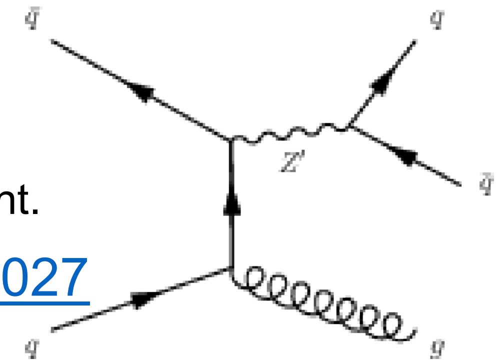


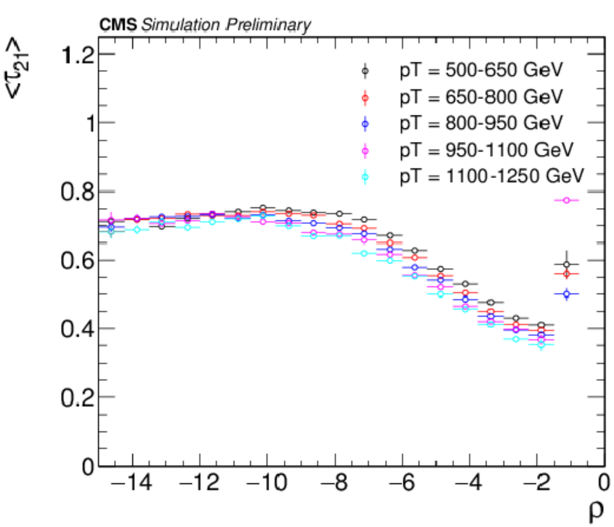
- Currently in preparatory stage with 13 TeV data
- Will require different strategy than bump hunt due to wide nature of signal.
  - Exact background prediction (normalization+shape) very important.

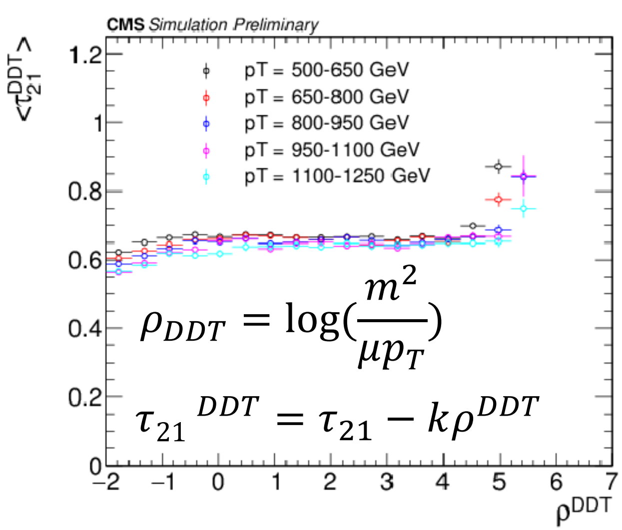
Not so high mass resonance

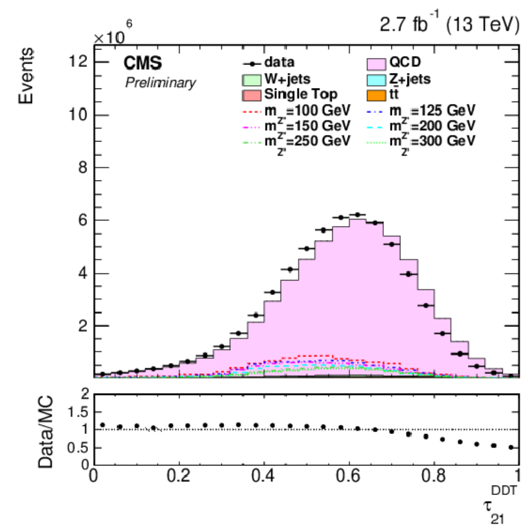
[CMS-EXO-16-030](CMS-EXO-16-030)

- A light mass Z'->qq resonance
- Z' with a recoil to get into boosted regime
  - Ease of triggering (trigger uses substructure too!)
  - Allows to go lower in mass than possible using a dijet bump hunt.
- Uses the designed decorrelated tagger (DDT) [1603.00027](1603.00027)
  - Takes out dependence of τ21 and jet mass
  - Uniform jet mass window across p_T range



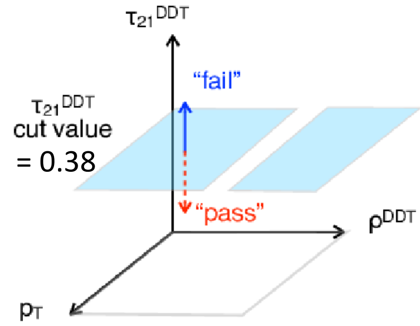




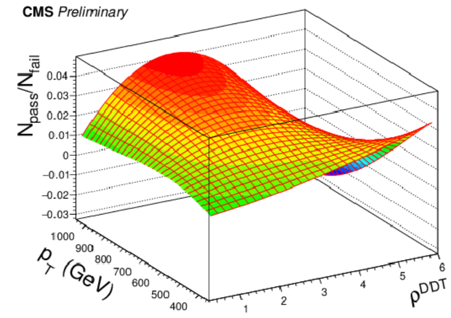


Not so high mass resonance

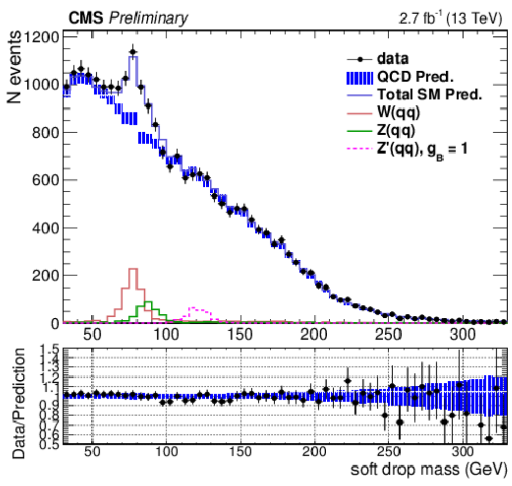
- Background estimated from the data
- Mass and τ_{21}^{DDT} sidebands used to predict multijets background.
- Fit of signal and background shapes used to compute limits on the signal cross section.

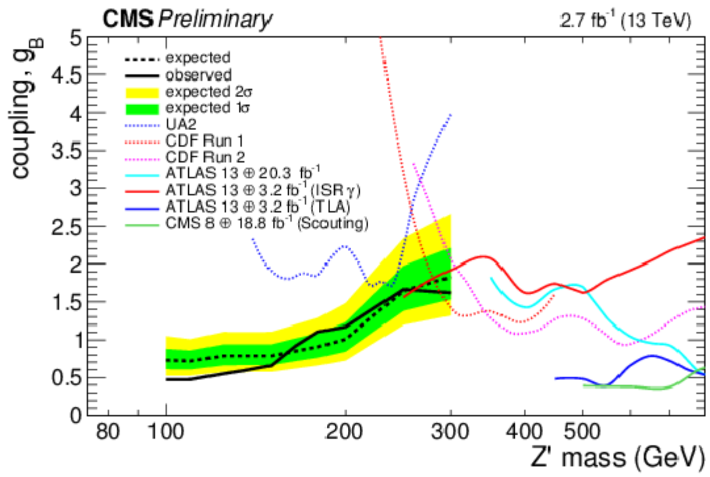


Sidebands



Pass/ fail ratio





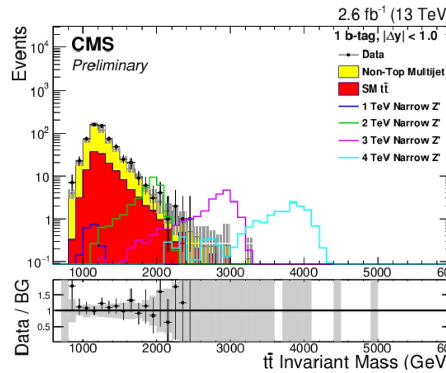
[CMS-EXO-16-030](CMS-EXO-16-030)

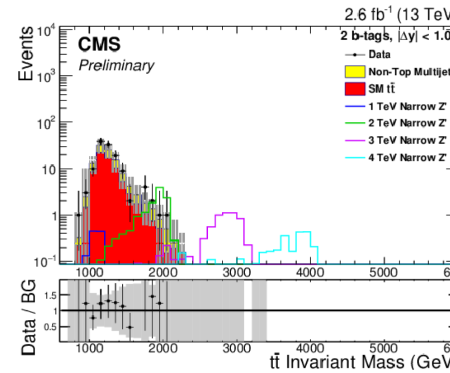
$t\bar{t}$ resonances

$t\bar{t}$ resonance searches

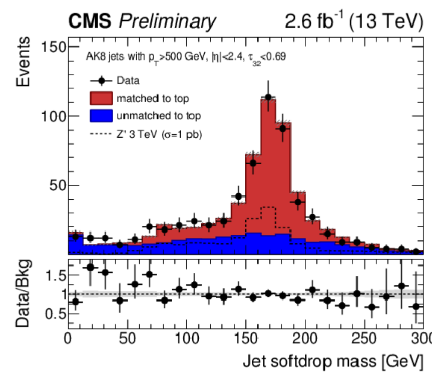
CMS-B2G-15-002, CMS-B2G-15-003

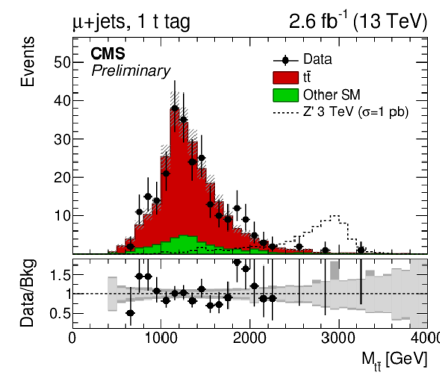
- Semi-leptonic and fully-hadronic final states.
- Hadronic analysis:
  - 2 top-tagged jets with event categorized using N(b jet) and $\Delta\eta(jj)$.
  - SM ttjets background from MC
  - Non-top multijets using the data. mistag rate measured using control samples with invertecd top tag: $\tau_{32} > 0.69$





- Semi-leptonic analysis:
  - Leptons (e, μ), E_T^{miss}, and jets.
  - Events categorized using N(top jets).
  - M(tt) reconstructed using the leptonic and hadronic top quark candidates
  - SM tt+jets and W+jets are the main backgrounds.

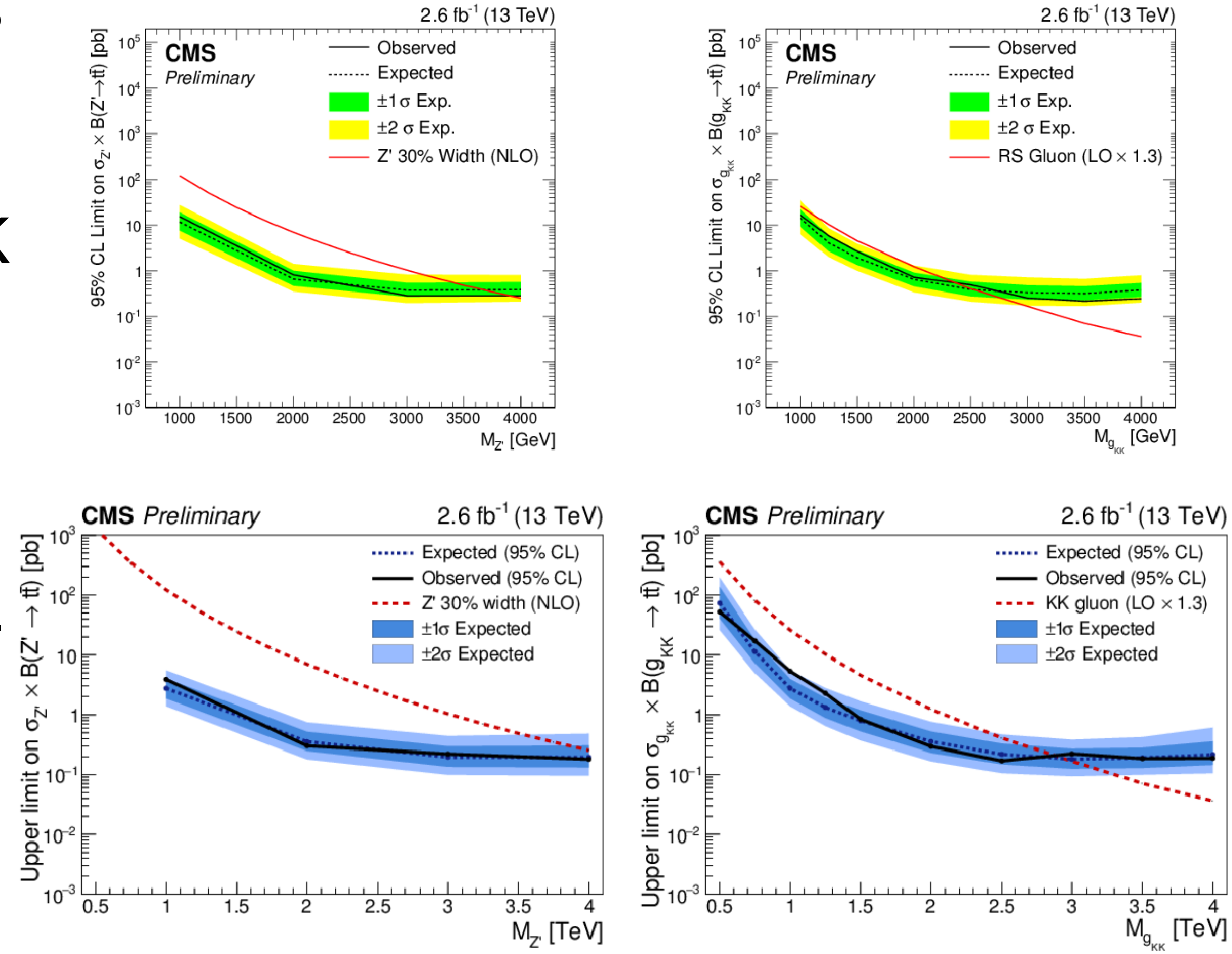




$t\bar{t}$ resonance limits

CMS-B2G-15-002, CMS-B2G-15-003

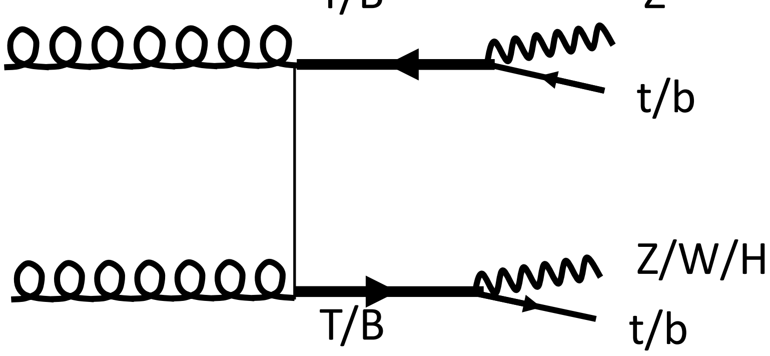
- No significant excess in the M(tt) spectra.
- Limits placed on a Z' decaying to tt and KK gluons of different widths.
- Comparable mass limits. Semi-leptonic better at low masses.
- Combination of the channels is in progress.



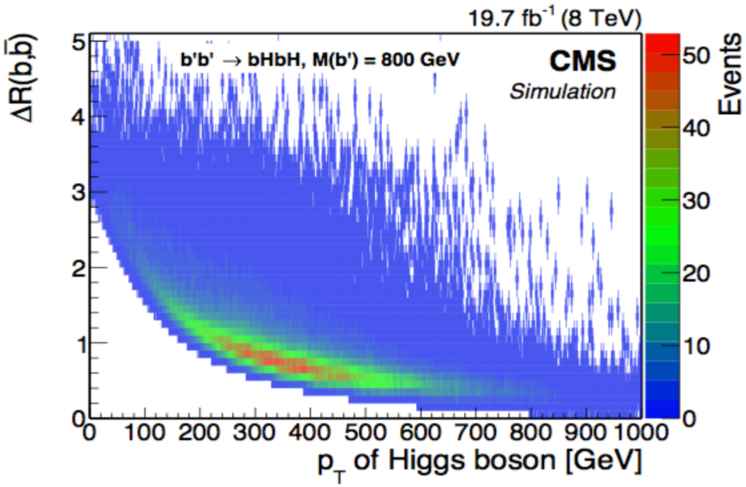
Vector-like quarks

Vector-like quark searches

- VLQs in BSM: Composite Higgs, GUT, extra dimensions
- Interesting properties:
  - Same transformation for left and right-handed particles => bare mass term respects gauge symmetry
  - Plays role of top quark partner to regularize the Higgs mass at higher orders
- Rich phenomenology: many production and decay modes

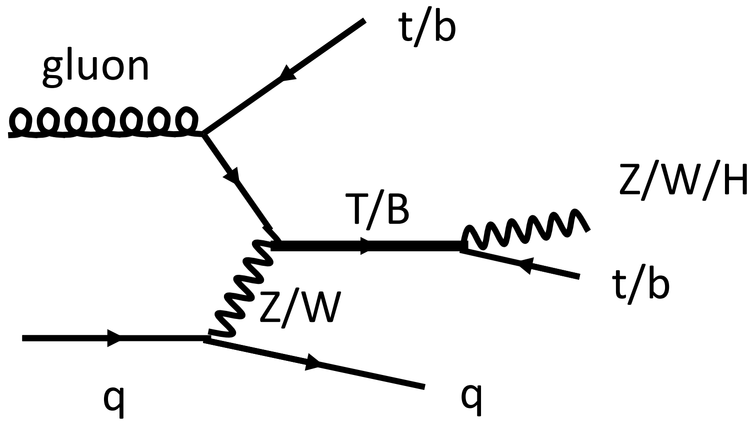


Pair-production: strong interaction

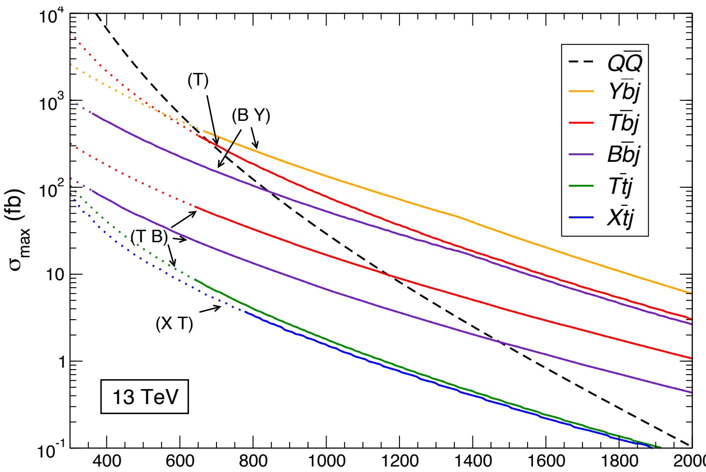


Single VLQ production the focus of Run 2

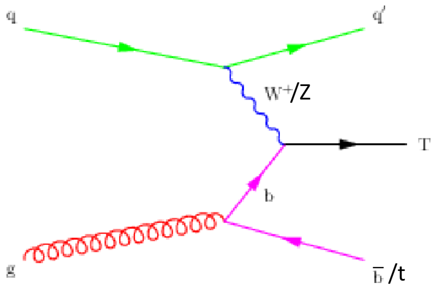
- Search for high mass vector-like top partners at 13 TeV
- Single production dominates in the high mass regime.
- Jet substructures heavily used
  - Trigger
  - Event reconstruction



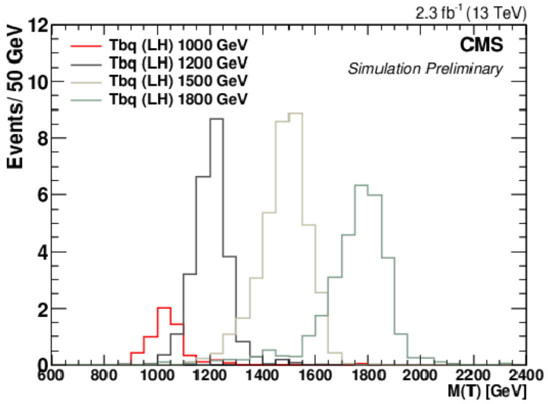
Single production: ewk interaction

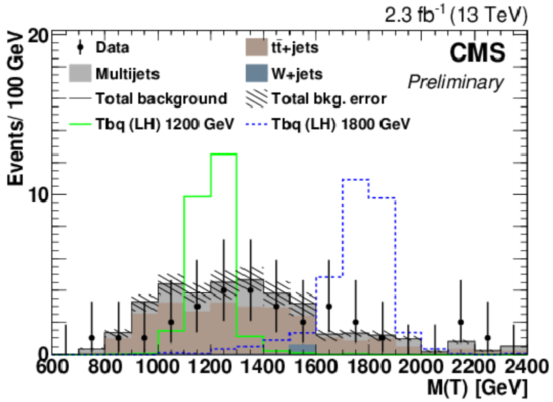


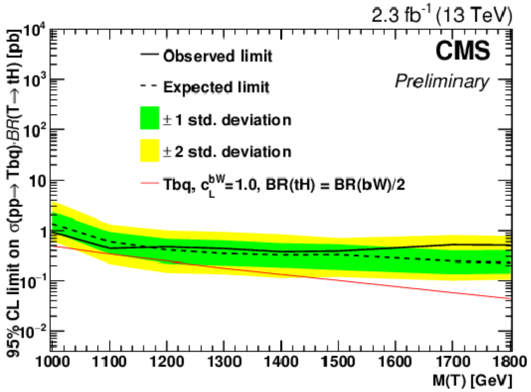
Vector-like top quark partner

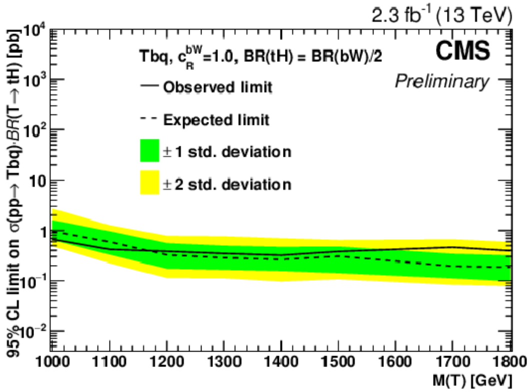


- Selecting top and Higgs jets
- Pairing them to find a resonance







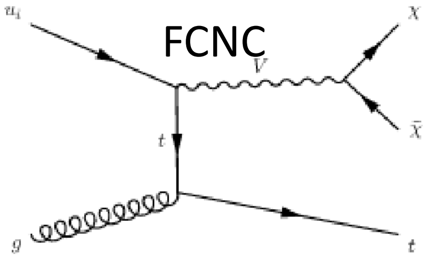


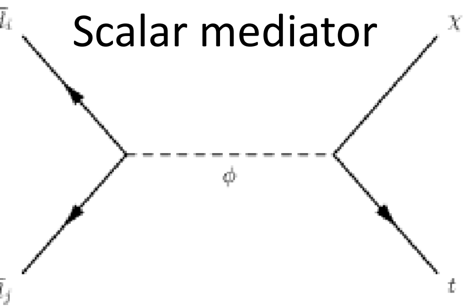
Very good signal resolution of ~5%

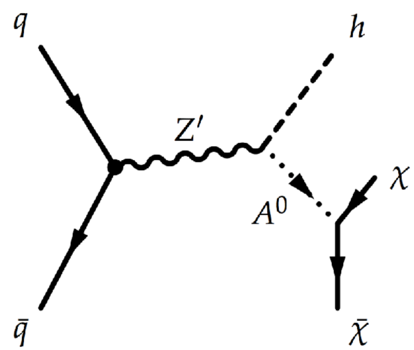
arXiv:1612.05336

Dark matter

Top and Higgs portal DM

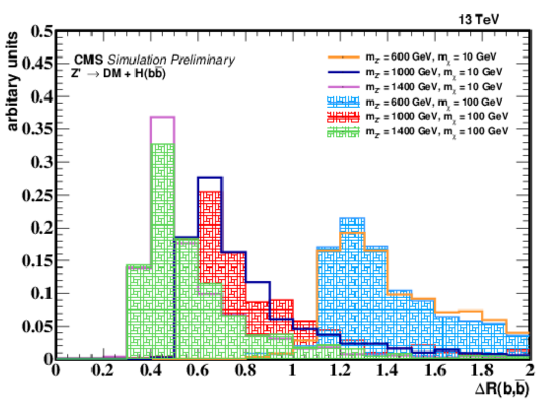




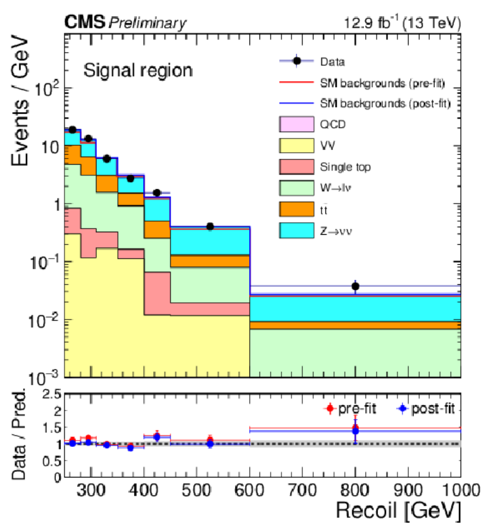


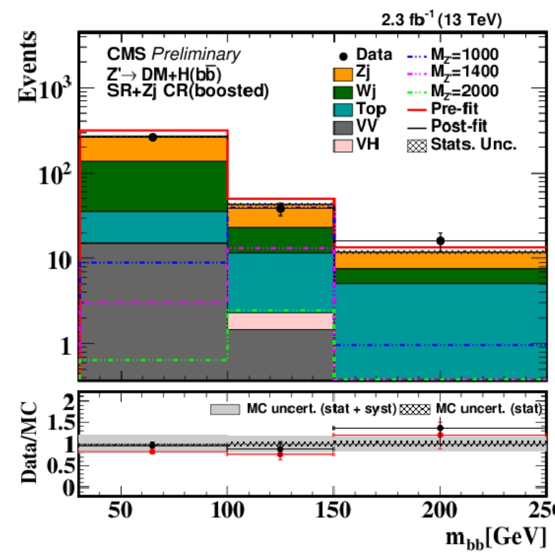
- Mon-top/ mono-Higgs signature of dark matter
  - Top tagging with AK8 jets
  - Higgs tagging using AK8 jets
  - Large missing E_T.

[CMS-EXO-16-012]



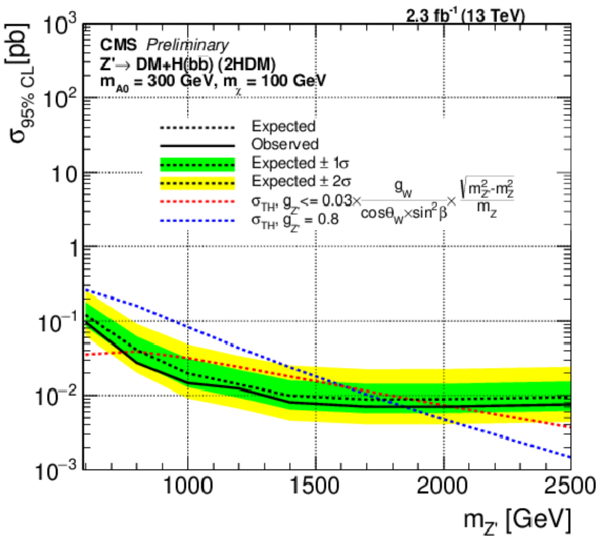
[CMS-CMS-EXO-16-040]

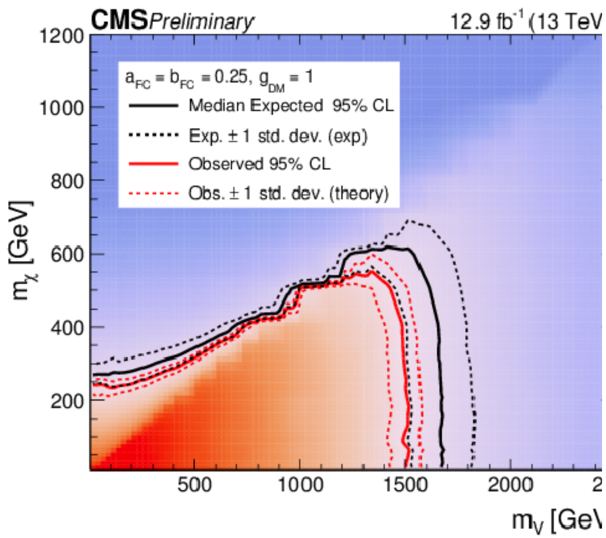


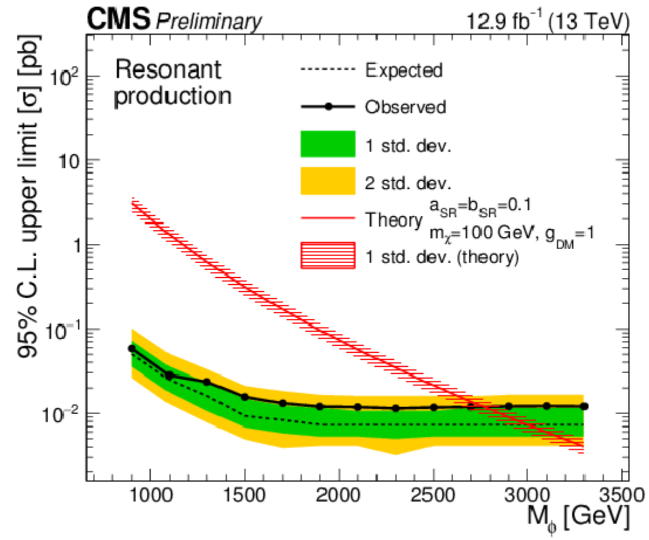


Mono-Higgs DM search

- Simultaneous fits in many control and signal regions yields limits on DM candidate and mediator masses



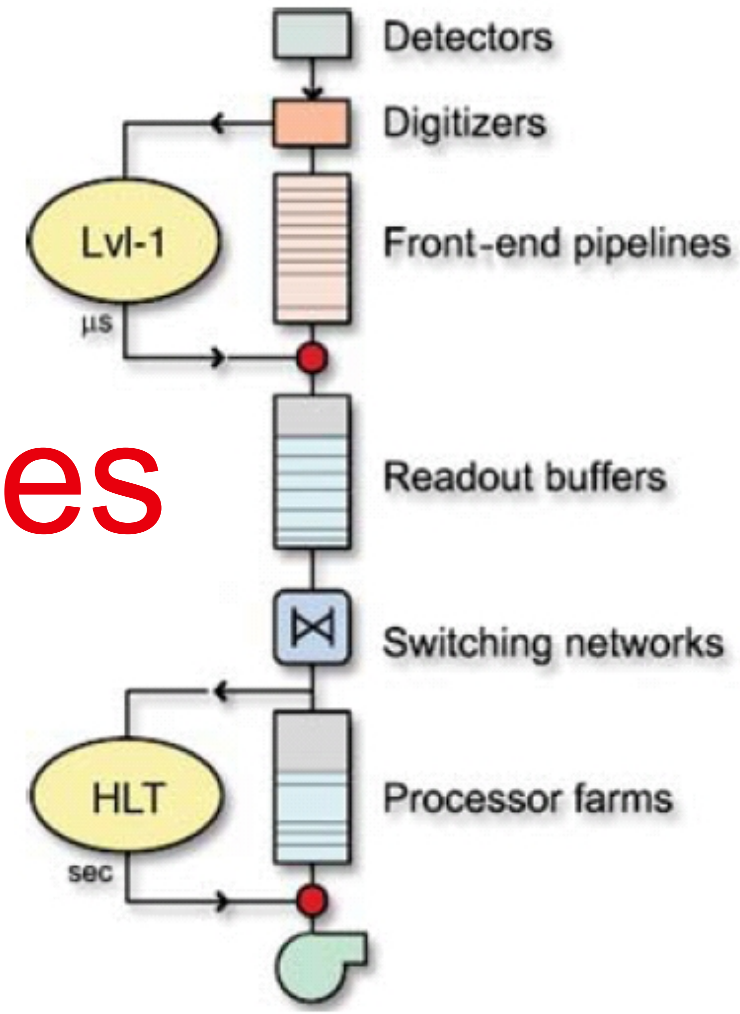




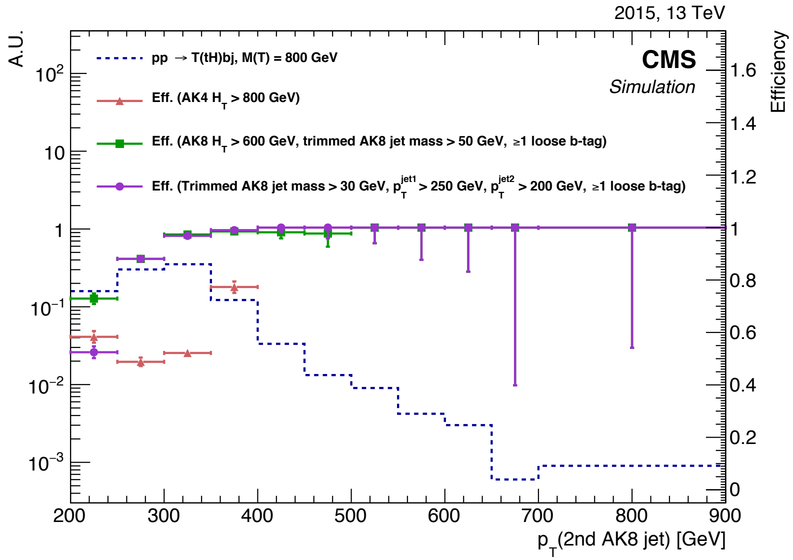
Trigger challenges

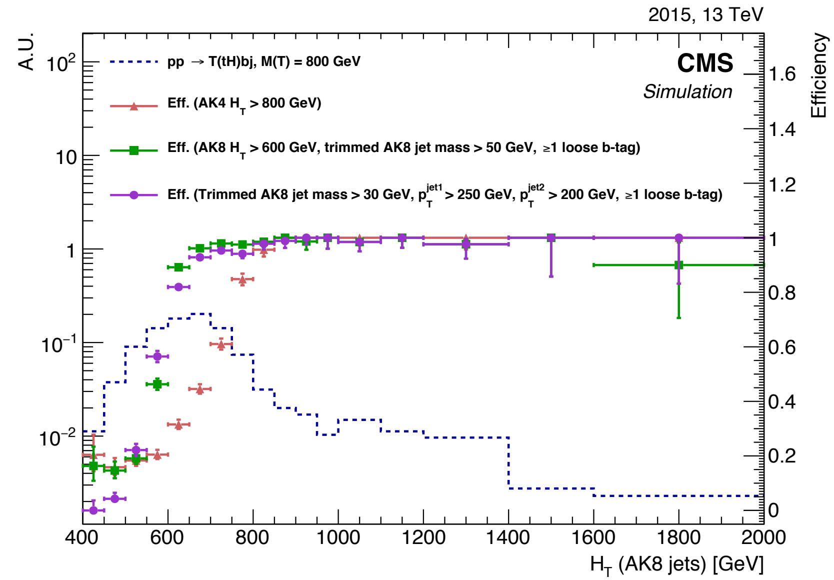
In all-hadronic environments

- Substructure to the rescue



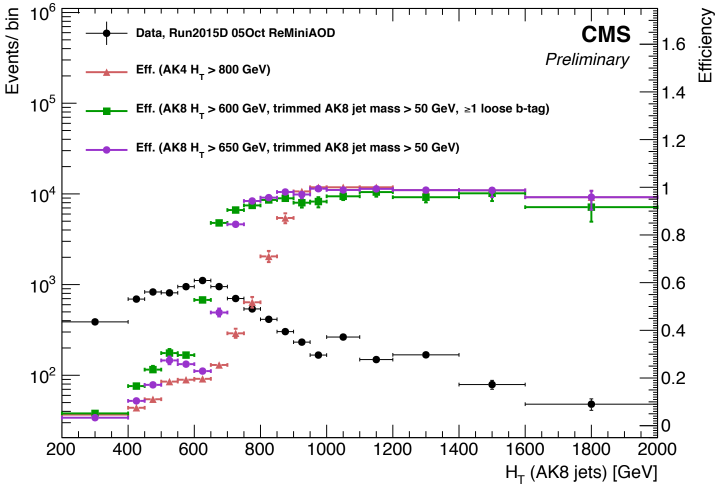
Triggering using boosted jets

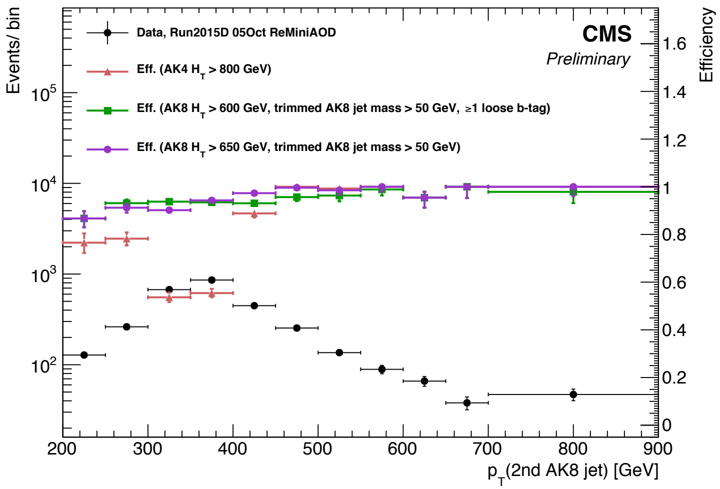




- Trigger utilizes features of an all-hadronic T->tH event:
  - Boosted top or Higgs: High pT jet with substructures and b-tagging
  - Overall large jet activity: large H_{T}
- Keeps trigger rates low while retaining high signal efficiency.

Performance in the data





- Performance in the data follows expectation from simulation studies.
- Deployed for collecting single vector-like top events since Summer 2015

Outlook

- Run 1 vs Run 2:
  - Discovery mode started with the Higgs boson
  - Run 2 is warming up: Holds a lot of promise for new physics searches.
- Lessons learnt from Run 1:
  - Theory: Excluded ranges aka "where not to look for signals".
  - Understanding the LHC machine and the detectors.
  - New reconstruction techniques.
  - Better analysis strategies.
- Current Run 2 goals:
  - Perfect new detection techniques like boosted boson and top tagging.
  - Search for signals in very high masses (out of reach at 7 and 8 TeV).
- LHC should deliver 100-200/fb of data at the end of Run 2:
  - Will be sensitive to a large variety of new physics signals.

An incomplete list – 2016 dataset

- (2016 ICHEP dataset) All-had VV resonances B2G-16-021
  - http://cds.cern.ch/record/2239381?ln=en

res

- (2016 ICHEP dataset) Search for VW diboson resonances in semi-leptonic final state B2G-16-020
  - http://cds.cern.ch/record/2205880?ln=en

res

- (2016 ICHEP dataset) Mono-V DM search EXO-16-037
  - https://cds.cern.ch/record/2205746/files/EXO-16-037-pas.pdf

DM

- (2016 ICHEP dataset) Mono-boosted top DM search EXO-16-040
  - https://cds.cern.ch/record/2205286

DM

An incomplete list– 2015 dataset

- 2015 data Search for ZV Resonance in Semi-leptonic Final States at 13 TeV B2G-16-010
  - http://cds.cern.ch/record/2199611?ln=en
- 2015 data HH->4b resonance B2G-16-008
  - http://cds.cern.ch/record/2202811?ln=en
- 8 TeV + 2015 13 TeV VV resonance B2G-16-007
  - http://cds.cern.ch/record/2154306?ln=en
- 2015 data Search for VH in the (l l, l nu, nu nu)bb final state B2G-16-003
  - http://arxiv.org/abs/1610.08066
- 2015 data Search for light vector resonances decaying to quarks at 13 TeV EXO-16-030
  - https://cds.cern.ch/record/2202715
- 2015 data all-hadronic T->tH B2G-16-005
  - https://arxiv.org/abs/1612.05336
- 2015 data Z'->tT all-hadronic B2G-16-013
  - http://cds.cern.ch/record/2217867?ln=en

res

res

res

res

res

VLQ

VLQ

An incomplete list– 2015 dataset

- 2015 data all-hadronic ttbar res. B2G-15-003
  - http://cds.cern.ch/record/2160237?ln=en

tt

- 2015 data semi-leptonic ttbar res. B2G-15-002
  - http://cds.cern.ch/record/2138345?ln=en

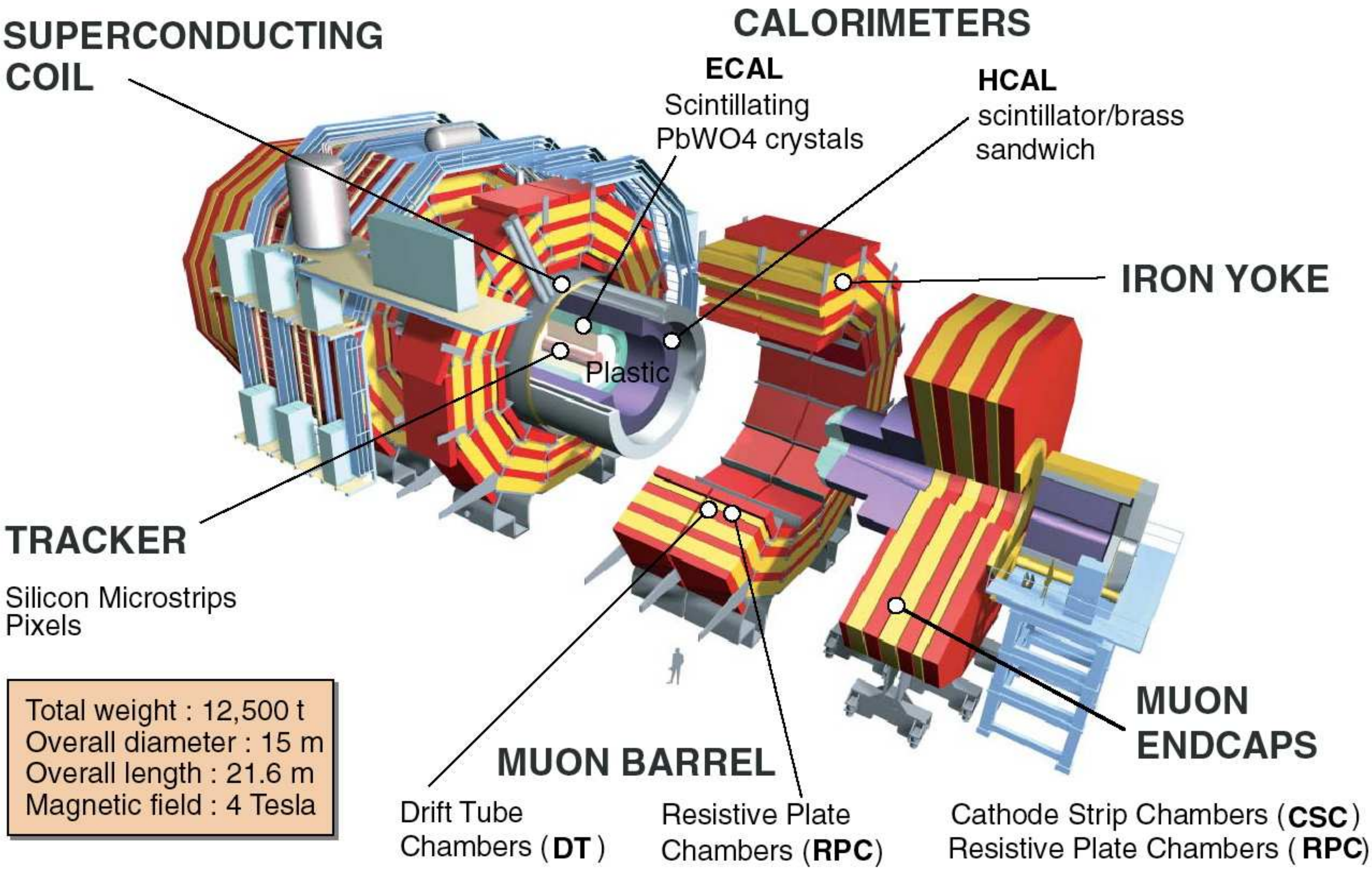
tt

- 2015 data Search for dark matter in association with a Higgs boson decaying into a pair of bottom quarks at sqrt(s) = 13 TeV with the CMS detector EXO-16-012
  - http://cds.cern.ch/record/2202804

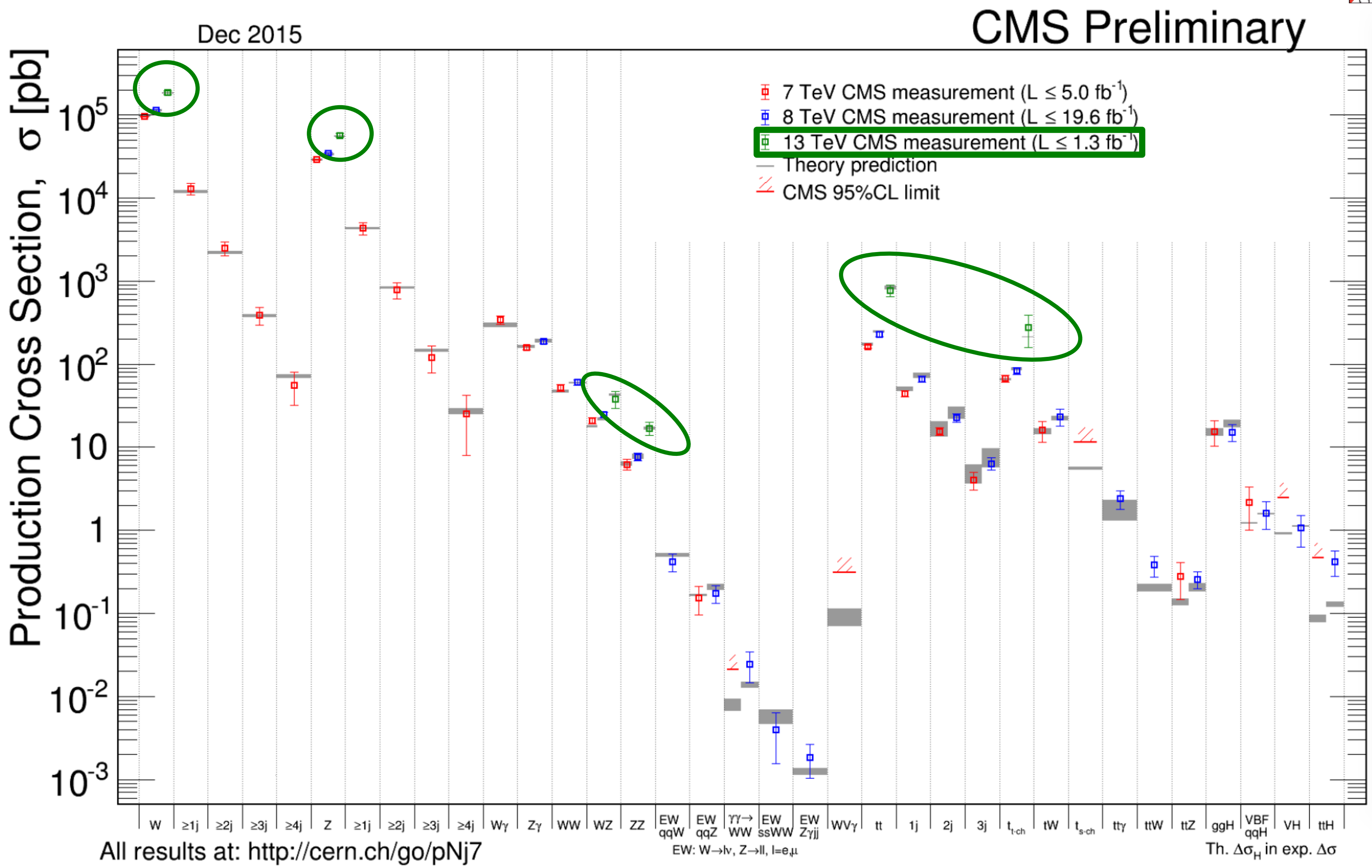
DM

BACKUP

The CMS detector



SM status



Dec 2015
CMS Preliminary
Production Cross Section, σ [pb]
7 TeV CMS measurement (L ≤ 5.0 fb⁻¹)
8 TeV CMS measurement (L ≤ 19.6 fb⁻¹)
13 TeV CMS measurement (L ≤ 1.3 fb⁻¹)
Theory prediction
CMS 95%CL limit
W ≥1j ≥2j ≥3j ≥4j Z ≥1j ≥2j ≥3j ≥4j Wγ Zγ WW WZ ZZ EW qqW EW qqZ γγ→WW EW ssWW EW Zγjj WVγ tt 1j 2j 3j t-ch tW s-ch ttγ ttW ttZ ggH VBF qqH VH ttH
EW: W→lν, Z→ll, l=e,μ
Th. Δσ_H in exp. Δσ
All results at: http://cern.ch/go/pNj7

Sequential jet clustering

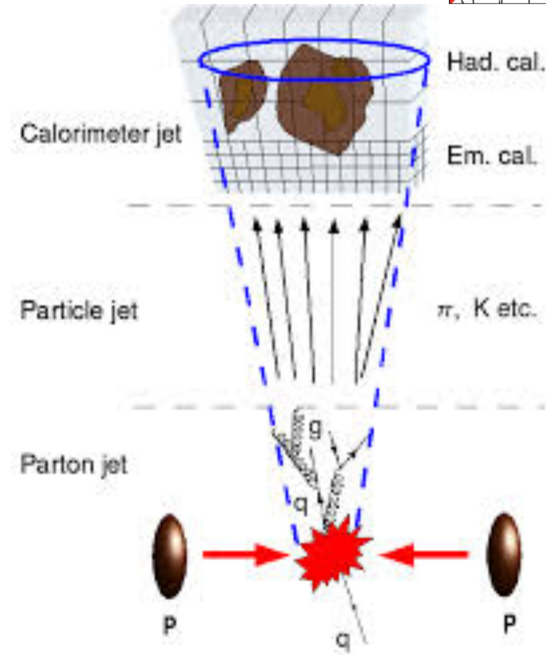
- Distance measure between two particles

$$d_{ij} = \min({k_{Ti}}^{2p}, kTj^{2p})\,\frac{\Delta_{ij}}{D}$$

  - $\Delta_{ij} = \sqrt{(y_i - yj)^2 + (\phi_i - \phi_j)^2}$
  - p = 0 for Cambridge-Aachen algorithm (CA)
  - p = 2 for kT algorithm
  - p = -2 for anti-kT algorithm
  - D = Distance measure of jets
- Find the smallest of d_{ij} and d_{iB}. Combine 4-momentum of particles *i* and *j*. B is beam direction. $d_{iB} = {k_{Ti}}^{2p}$
- Continue until all particles are clustered into jets.
- 2p = 0 for Cambridge-Aachen algorithm
- 2p = 2 for k_T algorithm
- 2p = -2 for anti-k_T algorithm

Jets

- Reconstructed charged particles (tracker+magnetic field) +Reconstructed neutral particles (calorimeters) +jet clustering algorithms = Jets



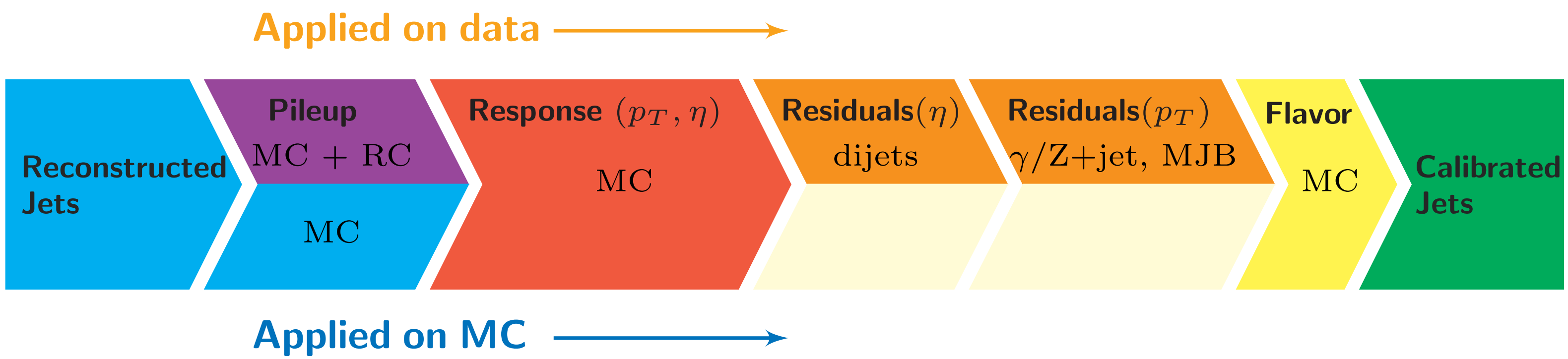
- CMS jet algorithms:
  - anti-kT (R-0.5)
  - Cambridge-Aachen (R=0.8, 0.15)
  - Anti-kT (R=0.4)

Run 1

Run 2

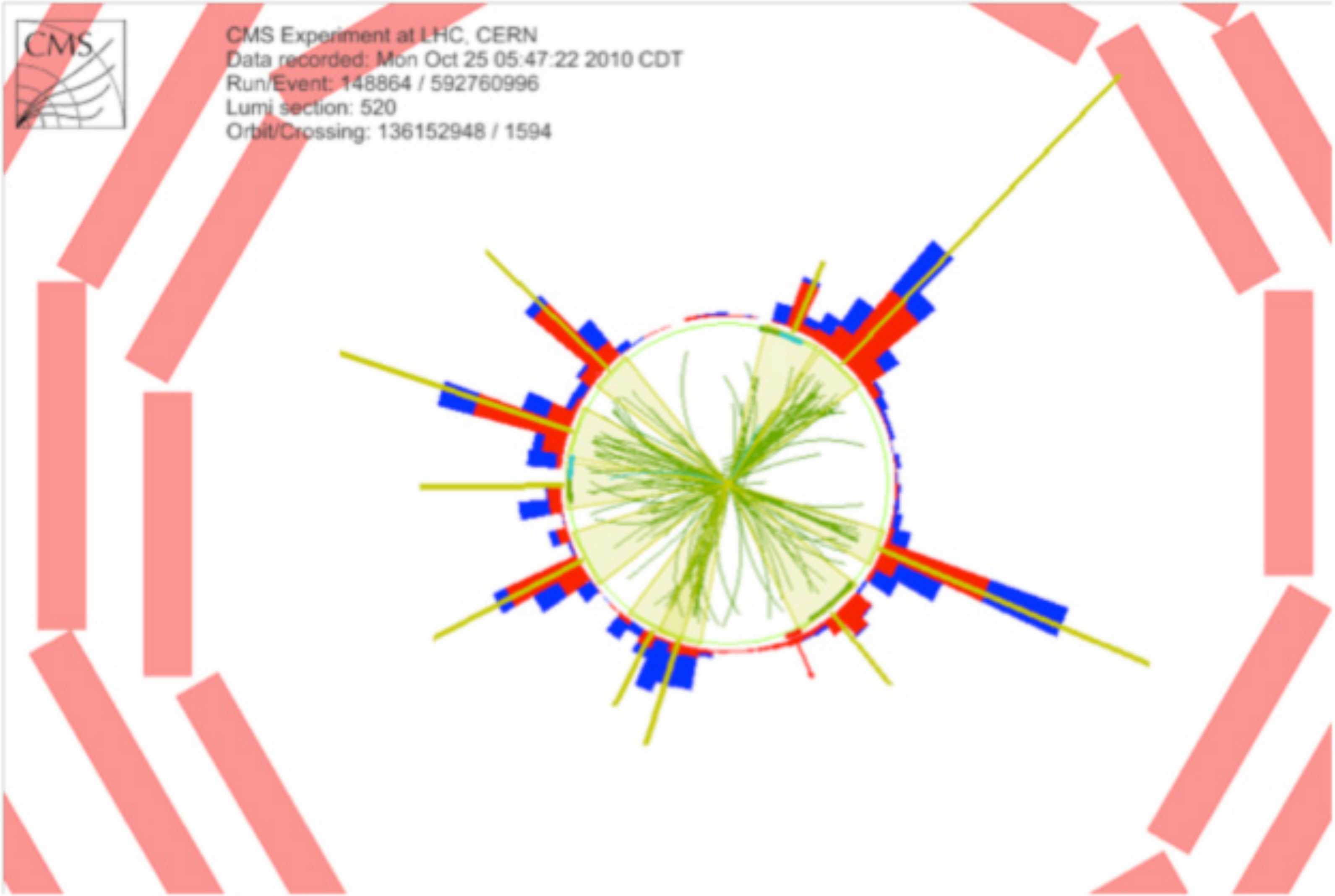
Distance parameter R gives the “cone” size of the jet.

CMS jet energy corrections

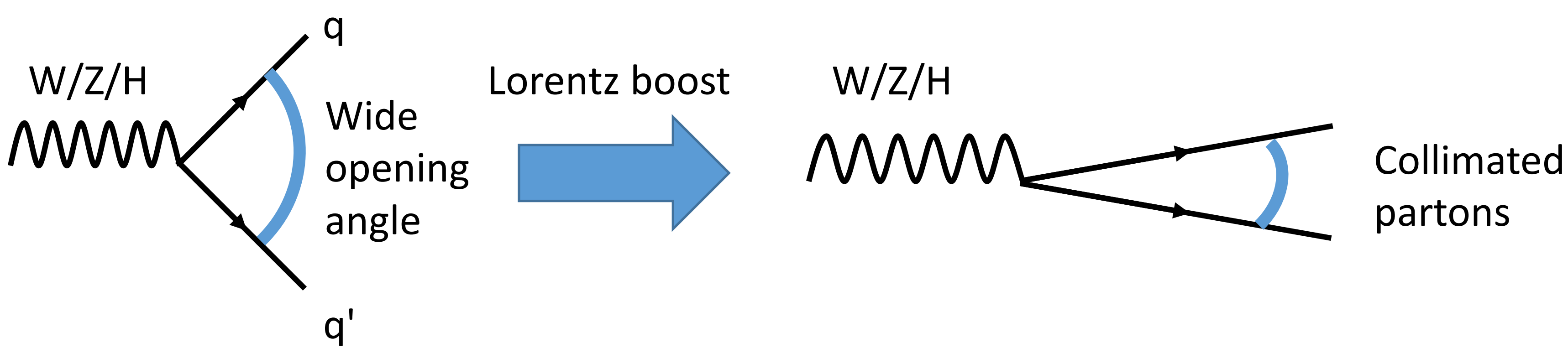


CMS-JME-13-004

CMS multijet event



Other kinds of jets

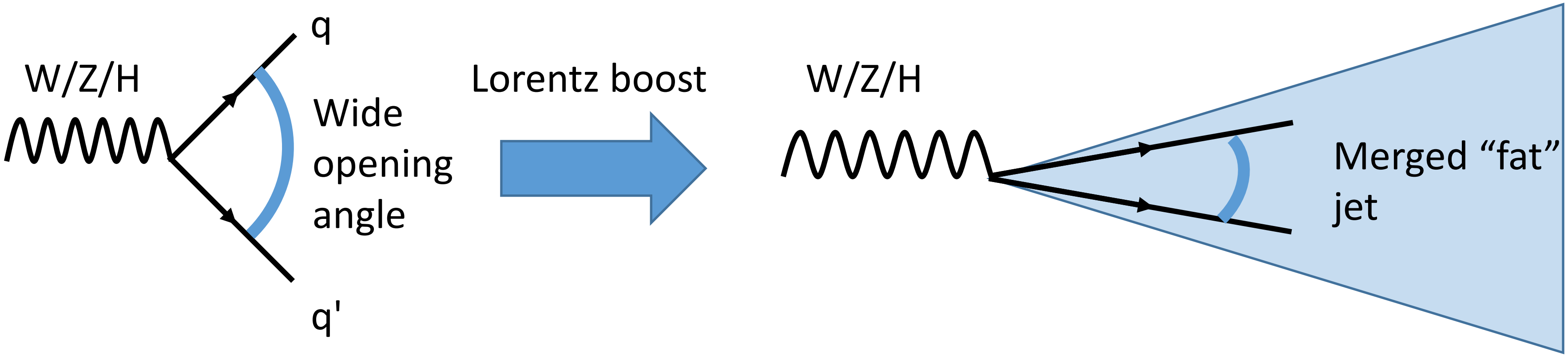


$$\Delta R = \sqrt{(y_1 - y_2)^2 + (\phi_1 - \phi_2)^2}$$

Rapidity

Azimuthal angle

Other kinds of jets



$$\Delta R = \sqrt{(y_1 - y_2)^2 + (\phi_1 - \phi_2)^2}$$

Rapidity   Azimuthal angle

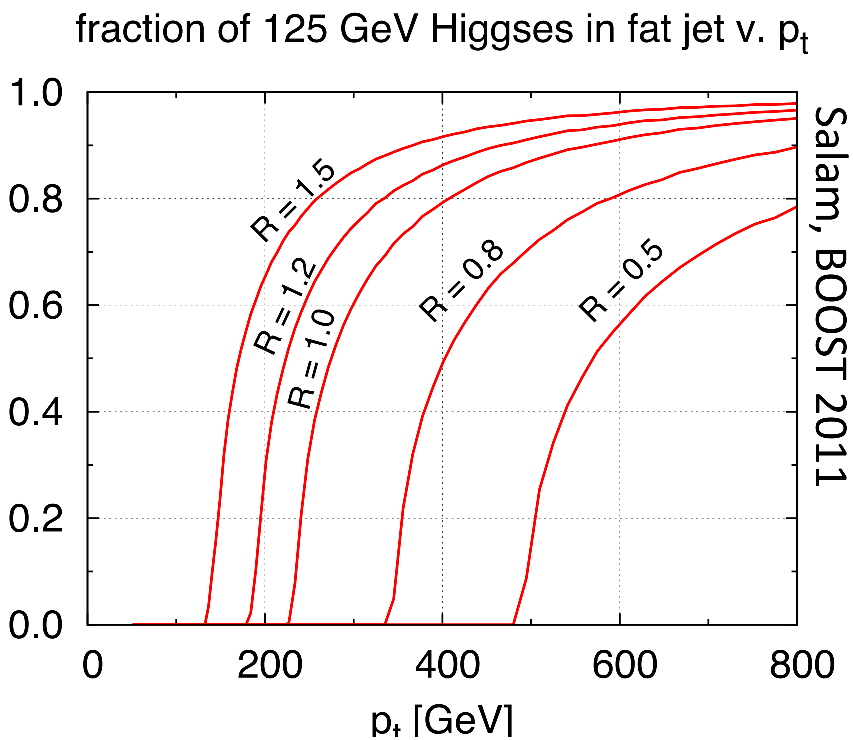
$$\Delta R \approx \frac{2M}{p_T}$$

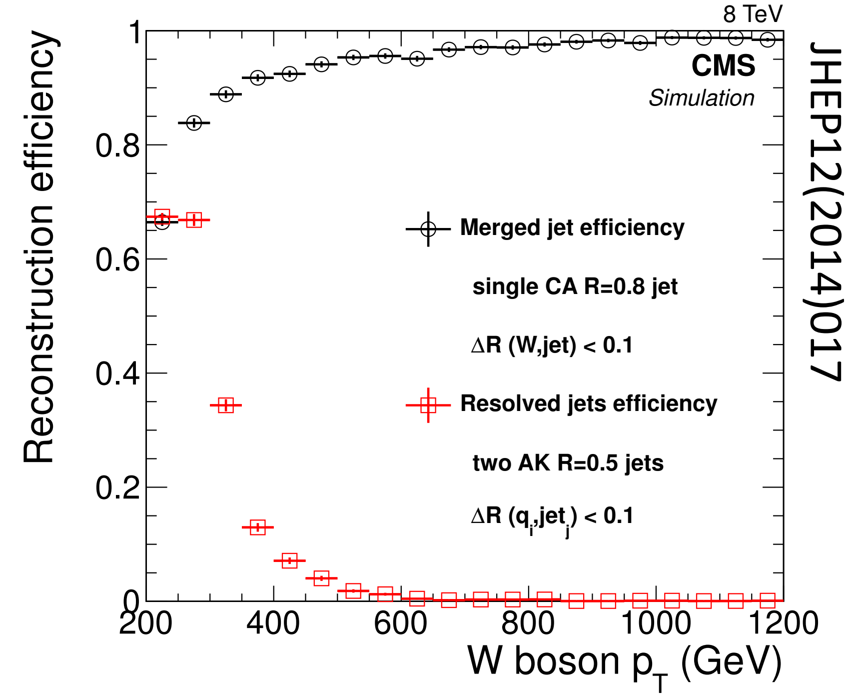
- Boosted SM particles can be reconstructed as a single jet.
  - Intrinsic jet mass
  - Internal substructure

Jet grooming

Or how to unravel subjets inside fat jets

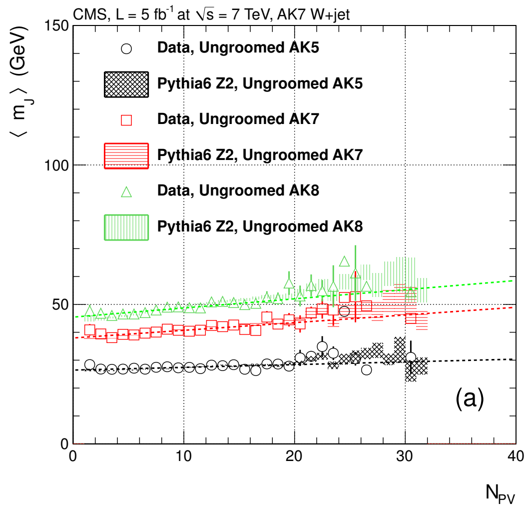
Why boosted jets



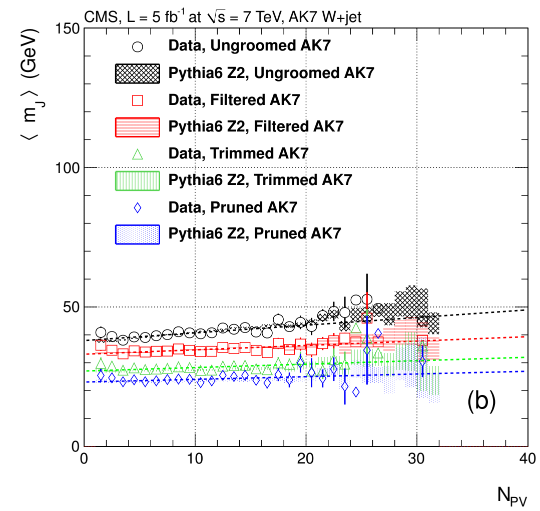


- Higher efficiency of reconstructing boosted W/Z/H or top quarks using fat jets than using slimmer resolved jets.

Effects of grooming



JHEP05(2013)090



JHEP05(2013)090

- Large mass acquired due to multiple pp interaction during the same bunch crossing *(pileup)*

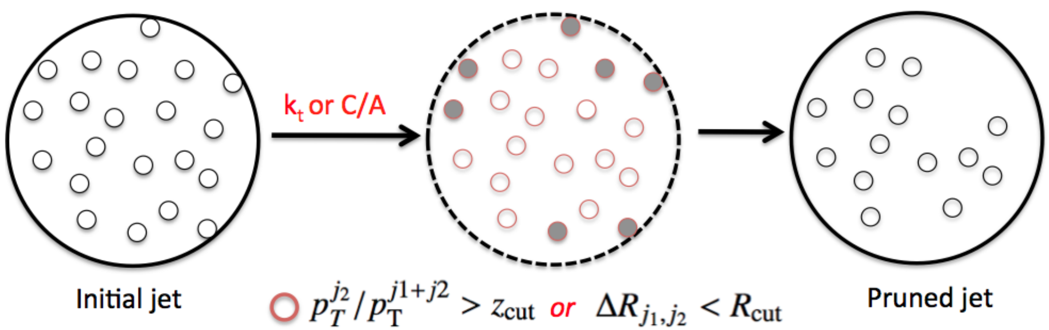
- Excess mass reduced by grooming

- Grooming reduces dependence of jet mass on pileup.

Pruning

- Removes the softest components of the jets.
- Jets reclustered using the Cambrdige-Aachen algorithm with a distance parameter of 0.8 (CA8 jets). In each recombination step a softer protojet is ignored, if it is soft enough or at a wide-enough angle in comparison to the original jet.

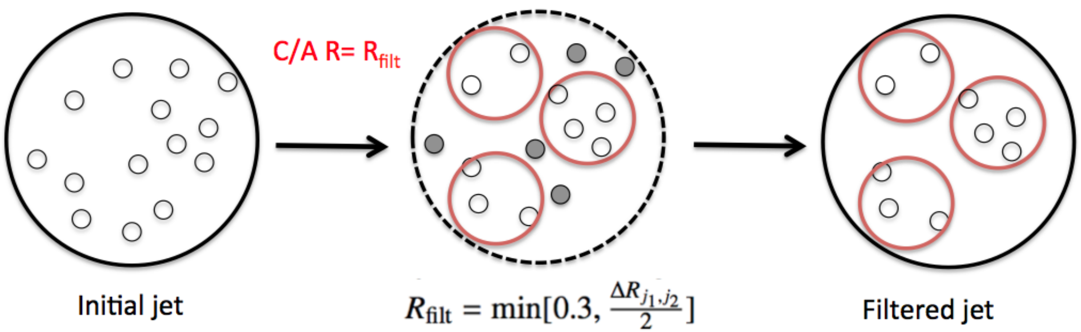
$R_{cut} < M(CA8)/p_T(CA8)$ $z_{cut} > 0.1$ arXiv:1306.4945



Jet filtering

- Fat jet constituents are reclustered using the Cambridge-Aachen algorithm with a smaller distance parameter (R = 0.3) compared to the parent jet. This defines new subjets, ordered in descending order of p_T.
- The four-momentum of the new jet is defined as the sum of the 4-momenta of the 3 hardest pT subjets.

arXiv:1306.4945



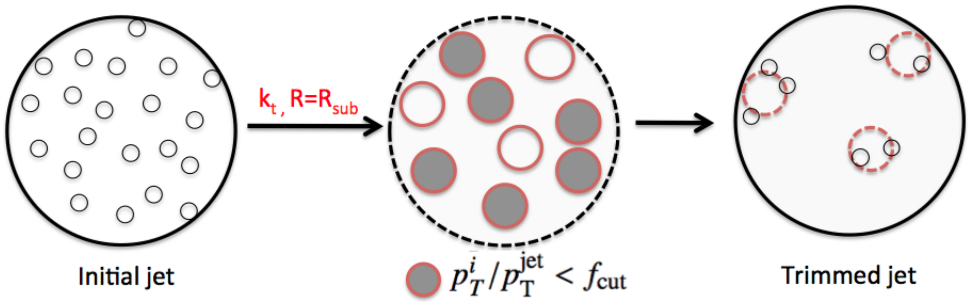
Jet trimming

- Reclusters jet constituents using the k_T algorithm with radius $R_{sub} = 0.2$
- If $p_T^{sub} > f_{cut}\lambda_{hard}$ keep, else remove.

Hard QCD scale = jet pT

Cutoff paramter = 0.03

arXiv:1306.4945



JHEP03(2011)015

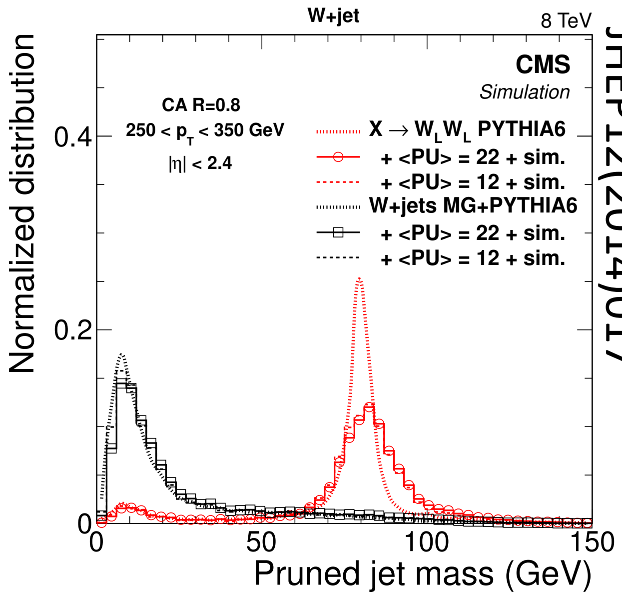
Jet N-subjettiness

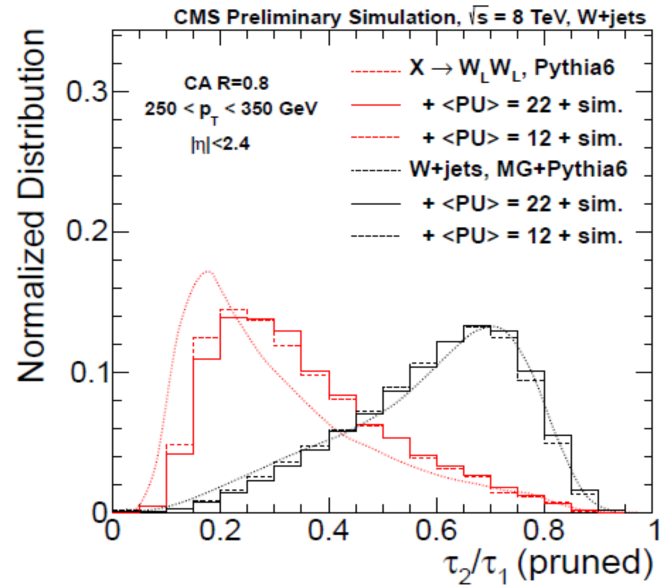
- Look for structrues inside jets by counting hard energy "lobes" inside the jet

$$\tau_N = \frac{1}{d_o} \sum_k p_{\mathrm{T},k} \min\{\Delta R_{1,k}, \Delta R_{2,k}, \ldots, \Delta R_{N,k}\}$$

- N = Number of candidate "subjets" without applying jet grooming techniques
- Subjet axes obtained using the exclusive k_{T} algorithm and reversing the last N clustering steps
- k = all jet constituents with transverse momentum $p_{\mathrm{T},k}$
- $\Delta R_{i,k} = \sqrt{(y_i - yk)^2 + (\phi_i - \phi_k)^2}$
- $d_o = \sum_k p_{\mathrm{T},k} R_o$, R_o = distance parameter of the original jet

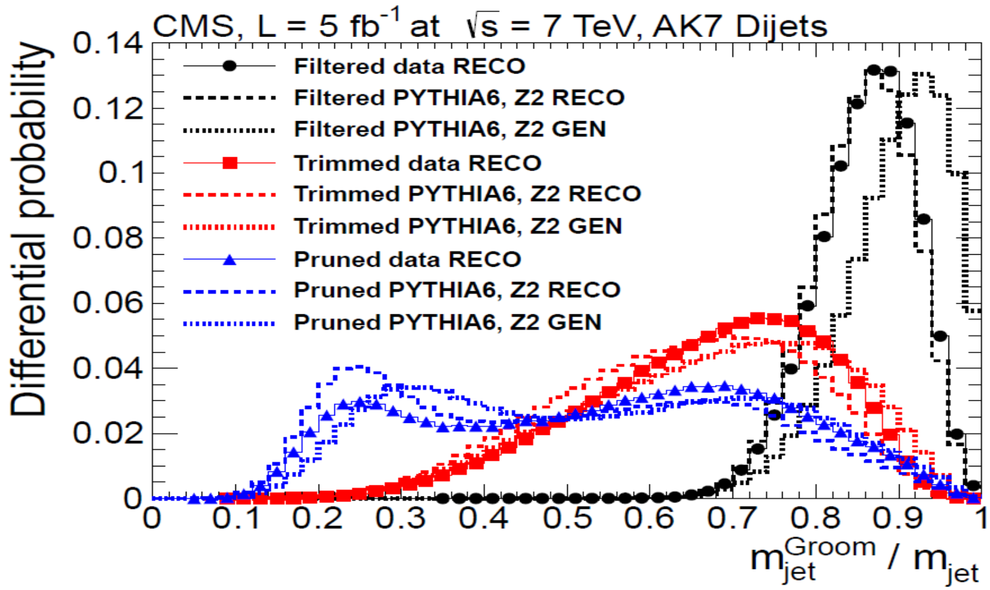
Jet pruning performance





- Pruning+N-subjettiness ratio τ_2/τ_1 gave the best performance for identifying boosted bosons in CMS in Run 1

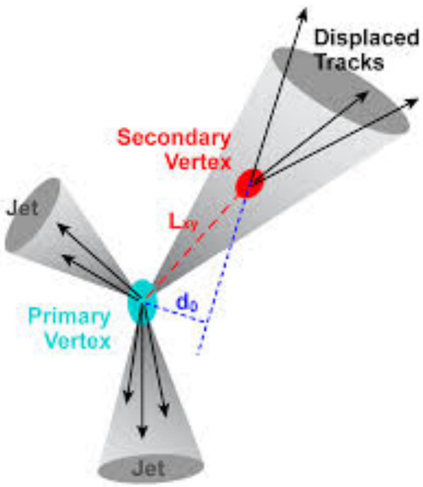
Comparing jet grooming techniques



JHEP05(2013)090

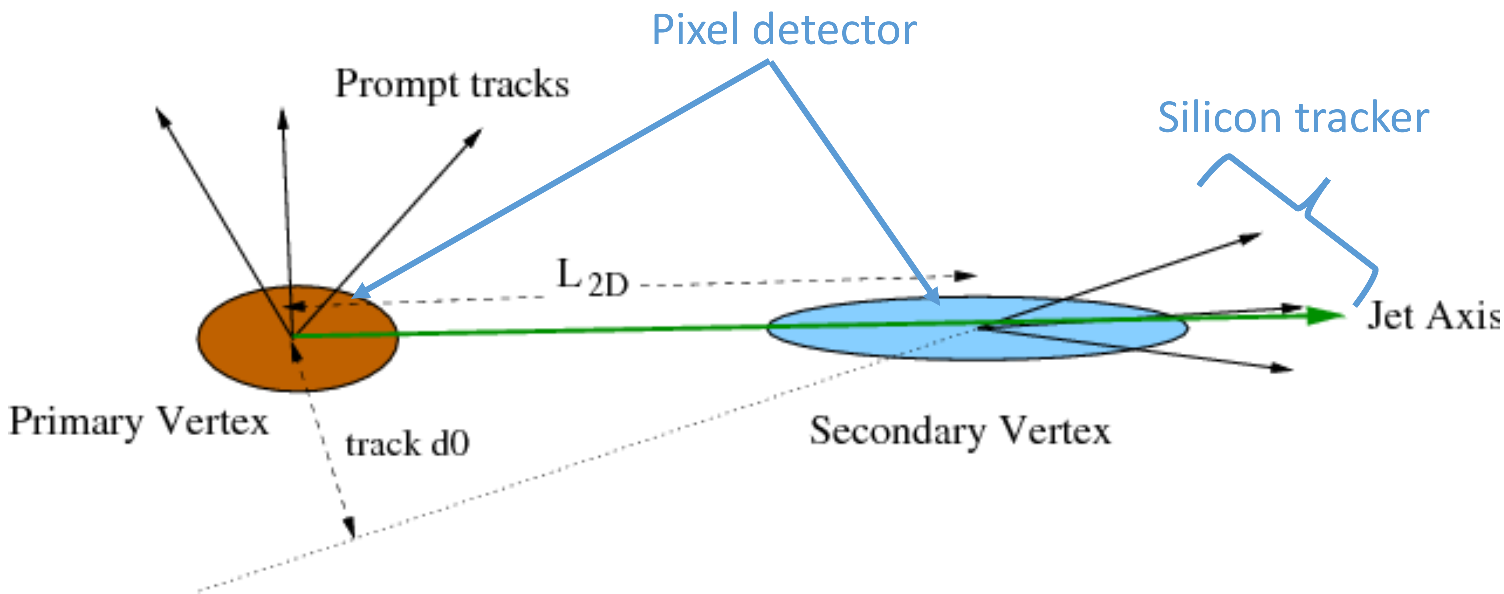
- Filtering is one of the least aggressive grooming algorithms, followed by trimming.
- CMS used filtered jets for top-tagging in Run 1.
- Trimming being used for triggering with jets and substructures.

B-tagging


Displaced
Tracks
Secondary
Vertex
Jet
Lxy
Primary
Vertex
d0
Jet

How b-tagging works

- Track and vertex info combined using a multivariate technique.
- Trained on simulated samples of jets of light and heavy flavours.
- Commissioned on collision data.
- Residual difference between simulation and data measured and accounted for as scale factors during data analysis.

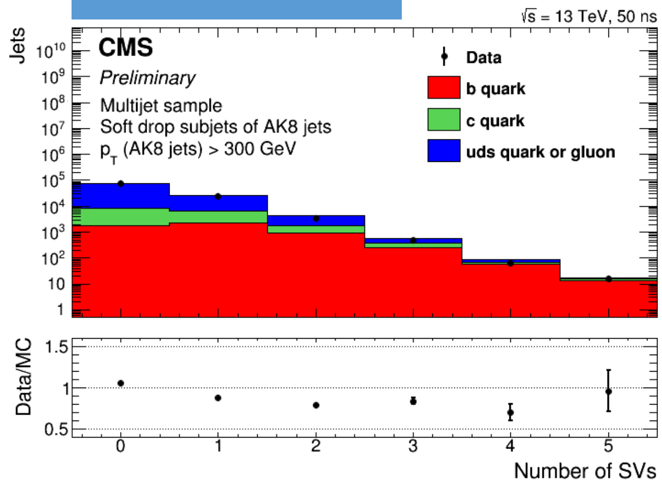


arXiv:1106.2516 [hep-ex]

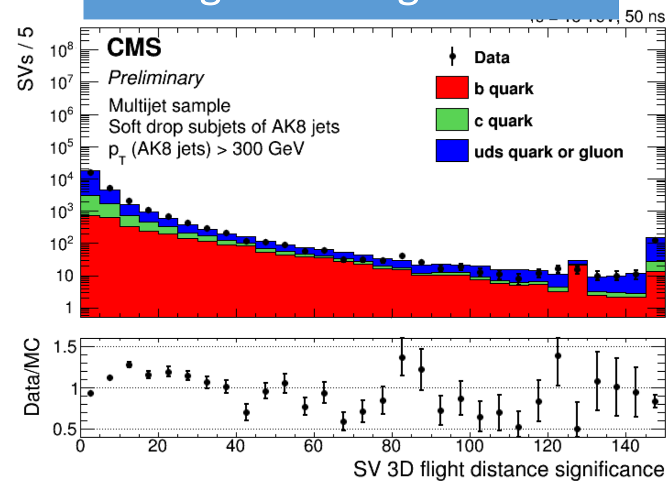
BTV13TeV50ns

Validation in the data

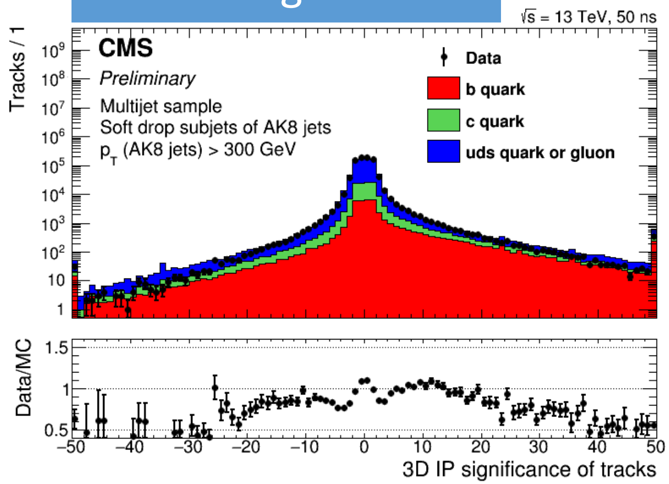
Number of SVs



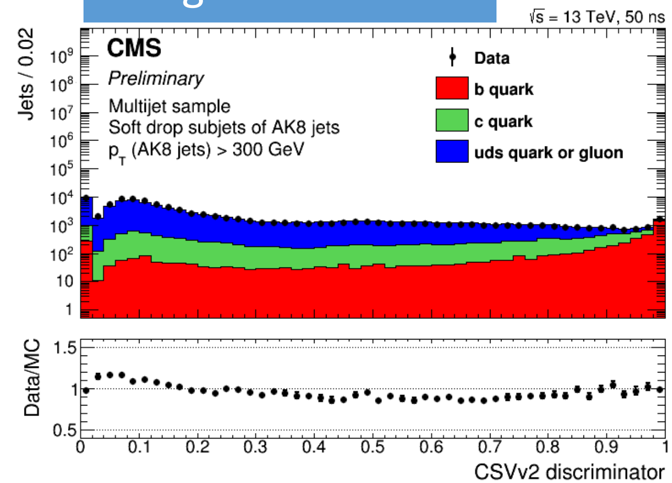
SV flight dist. significance



Track IP significance



B-tag discriminator



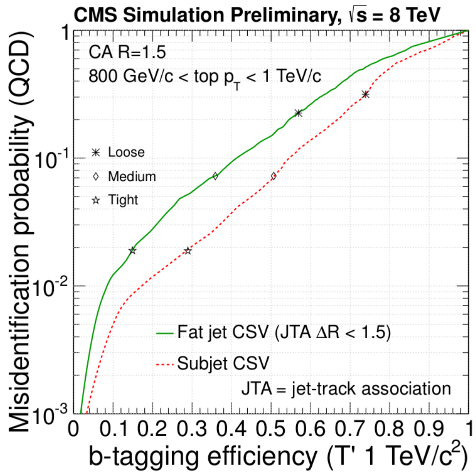
CMS-BTV-13-001

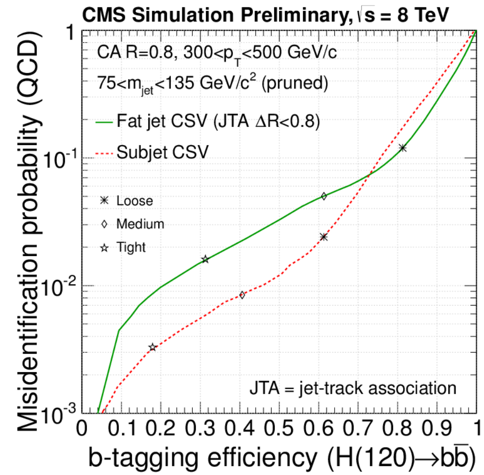
Performance

Top jets

Higgs jets

ROC curves





Tagging rate vs p_T