Searches with boosted topologies at LHC

Patrizia Azzi
INFN Padova (I)
for the ATLAS and CMS Collaborations

Rencontres des Moriond EWK 2014

Entering the boosted regime

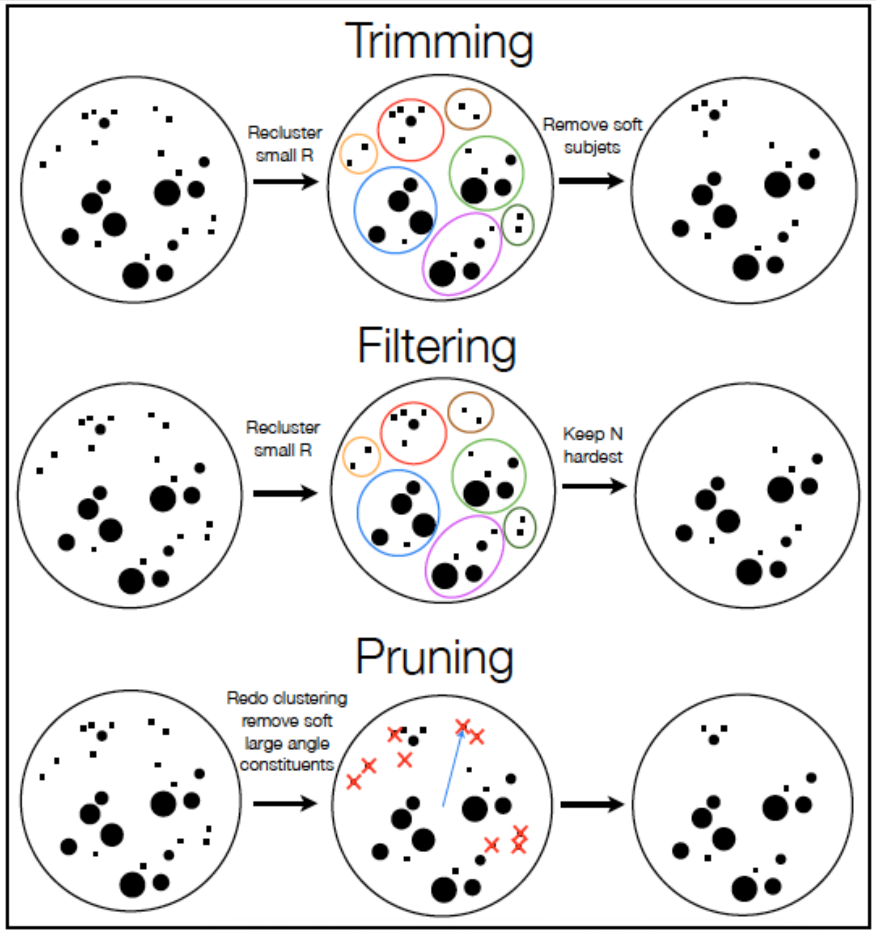
- The term « boosted » applies to particles with energy equal or above twice their mass: at those energies the opening angle of the decay products is about DR~2m/E
  - In the particular case of hadronic decays of W/Z/H/top the decay products end up collimated in a single jet
- New tools can be developed to profit of this new kinematic regime and extend the discovery potential
- Actually these ideas have been around for quite some time in the context of QCD studies, now getting a new life in the spotlight!

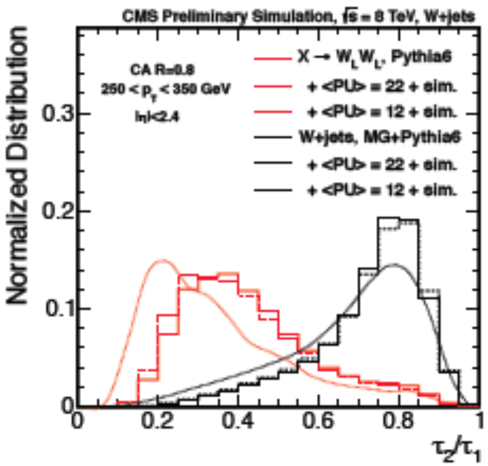
When? Why?

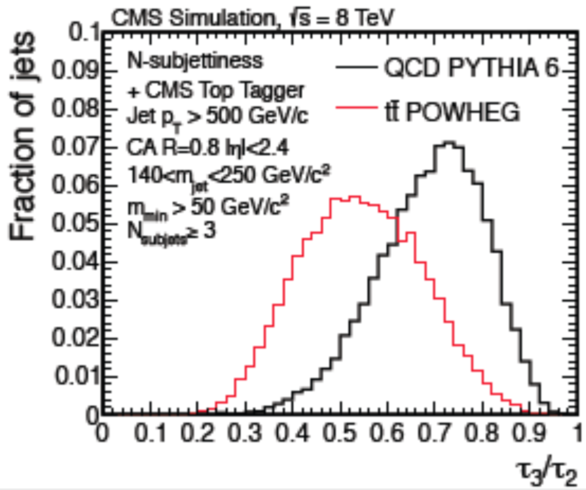
- **Decay of very massive BSM particles into object with a large mass difference**
- **Searching for phenomena on the tails of the pT distribution**
- **Important for both SM and BSM processes!**

Jet grooming & n-subjettiness

- <u>Jet mass is a fundamental parameter</u>
- Need « grooming » of the fat jet to remove:
  - unwanted soft QCD contributions
  - Pile Up
- Exploit the jet substructure and shape. <u>Many possible choices of discriminators</u>
  - <u>Example</u>: **n-subjettiness τ_N:** jet shape variable that express the consistency of the jet with having N or fewer sub-jets.
  - Better discrimination with ratios: τ_3/τ_2 for top or τ_2/τ_1 for W/Z





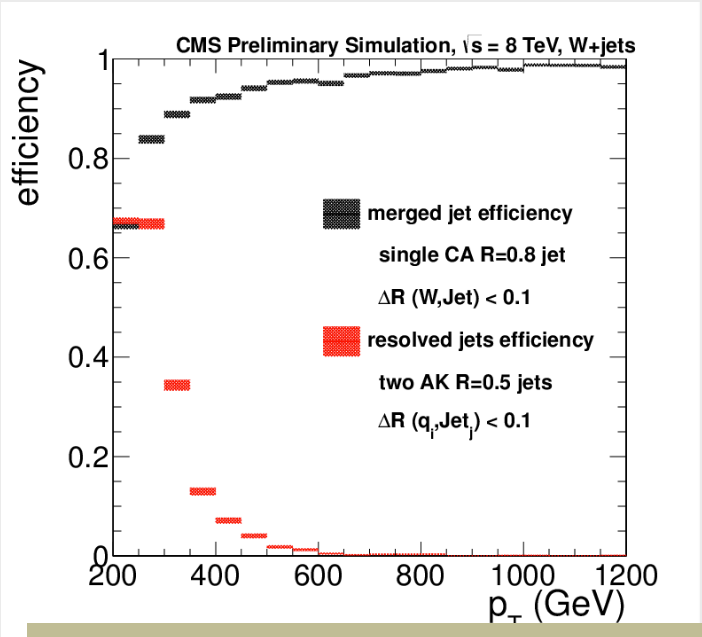


J.Thaler, K. Van Tilburg arXiv:1011.2268

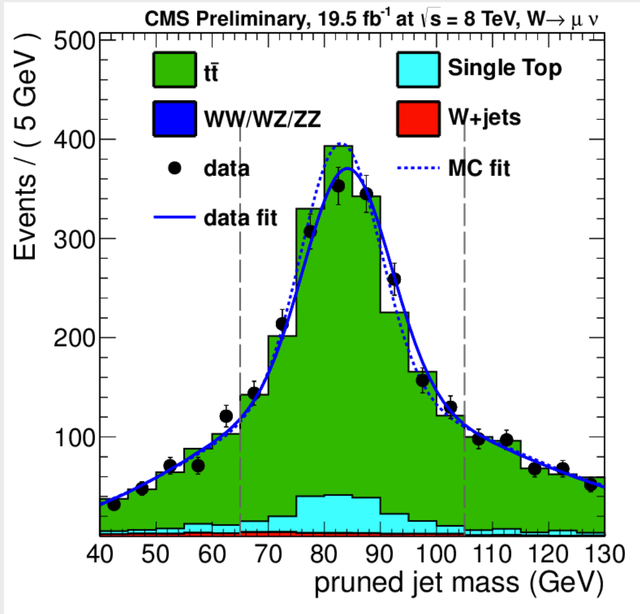
CMS-PAS-JME-13-006

ATLAS : JHEP 1309 (2013) 076

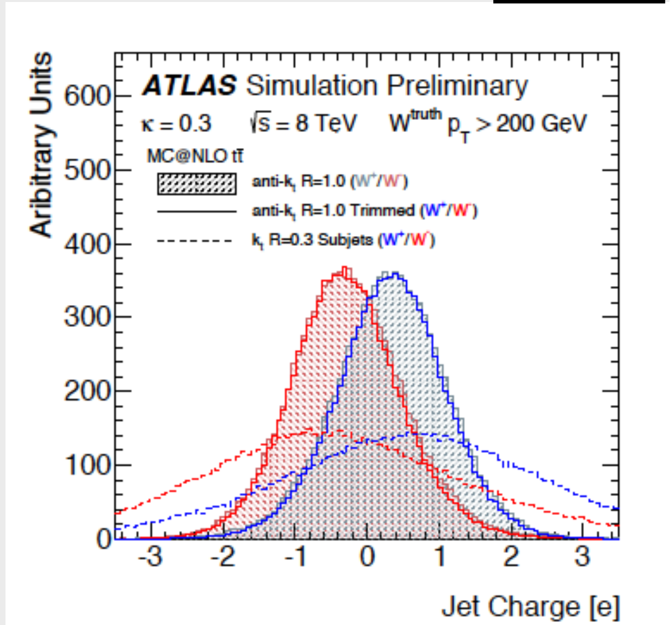
Boosted W/Z tagging



Efficiency for had W reco as 1 or 2 jets vs $p_T(W)$



W mass from T&P in tt events



Jet charge in W-tagged jets

- **For a W boson with $p_T \geq 200$GeV a $\Delta R=1.0$ jet captures most of the decay products.**
- W/Z boson ID is obtained taking a large jet cone (0.8-1.0) and performing:
  - cut on the groomed mass 60<M(W)<100 GeV
  - a cut on the likelihood of two-prongs in the jet (CMS uses n-subjettiness)
  - Also extends the concept of charge tagging to boosted Ws
- Validation from simulation AND data selecting W in tt events:
  - tt in l+jets, with btag + one « fat » high mass jet in the hemisphere opposite the lepton

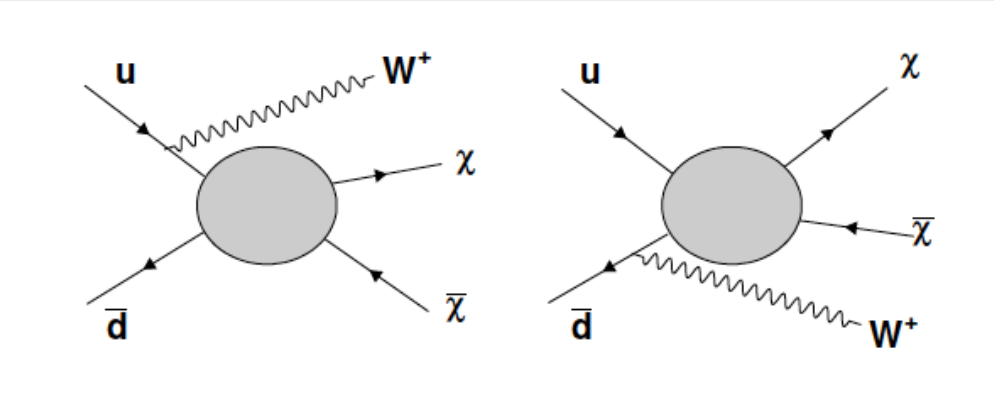
CMS-PAS-JME-13-006

ATLAS-CONF-2013-086

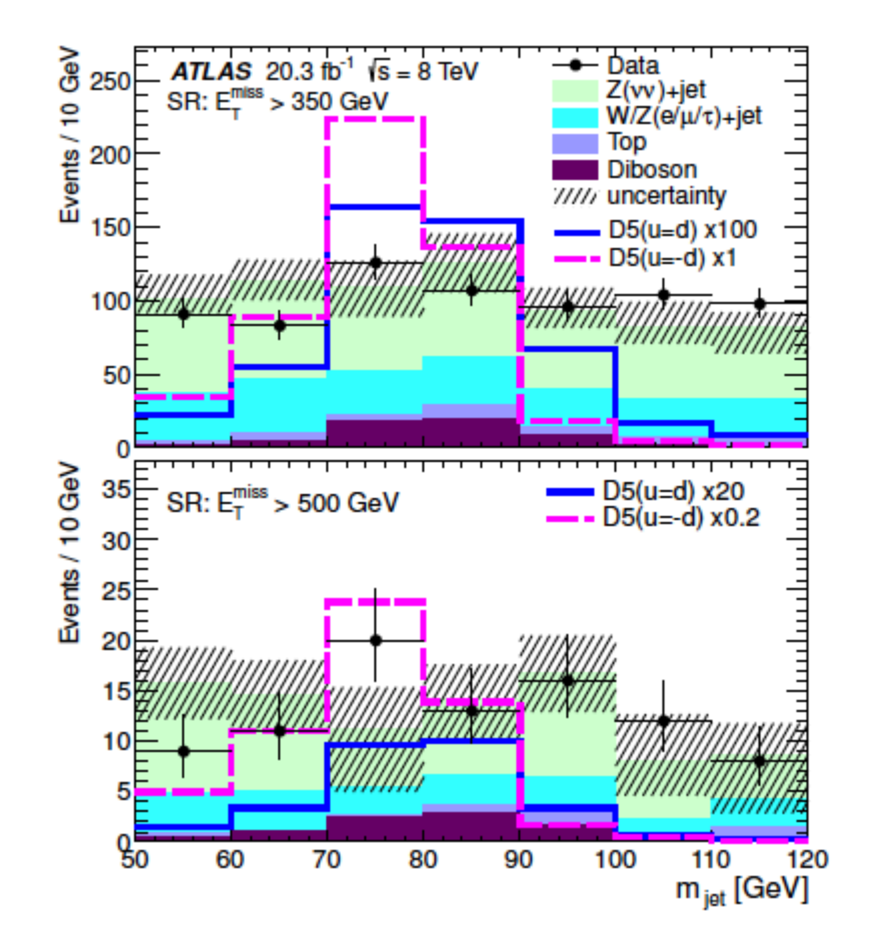
ATLAS-CONF-2013-087

Dark Matter search with boosted W

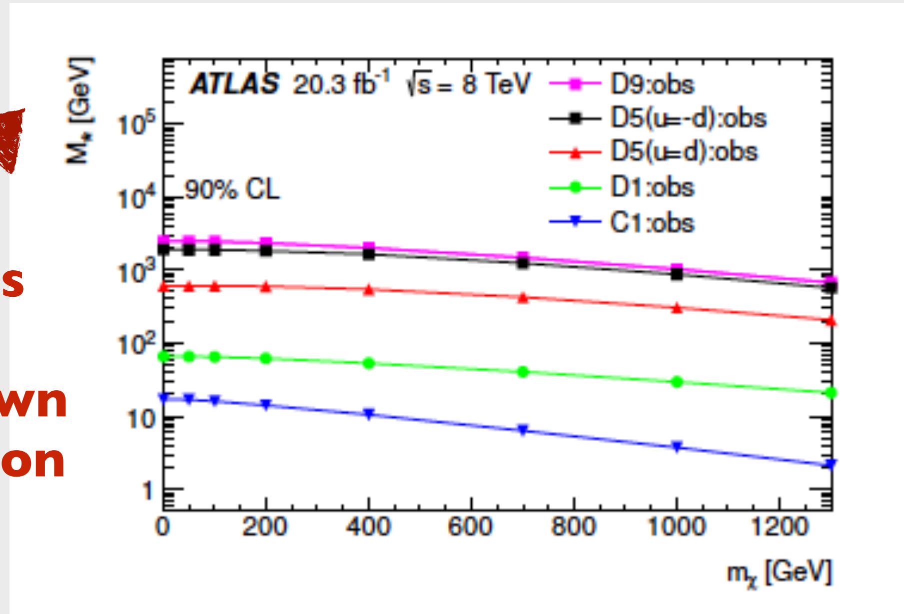
ATLAS PRL 112, 041802 (2014)



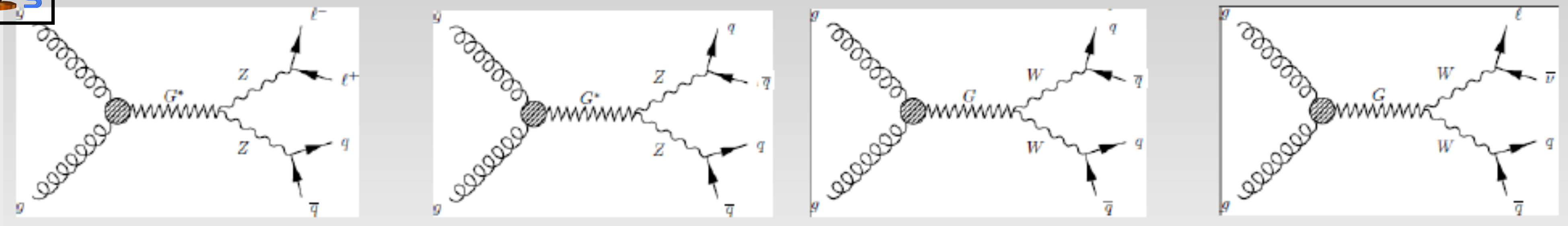
- W-Jet reco as CA=1.2 with mass drop and substructure checks on the two leading subjets:
  - **pT(jet)>250 GeV, Iηl<1.2,**
  - **50<M(jet)<120 GeV**
  - **Signal region:**
    - **E_T(miss)>350/500 GeV**
    - **veto extra narrow jets**



M* mass scale of unknown interaction



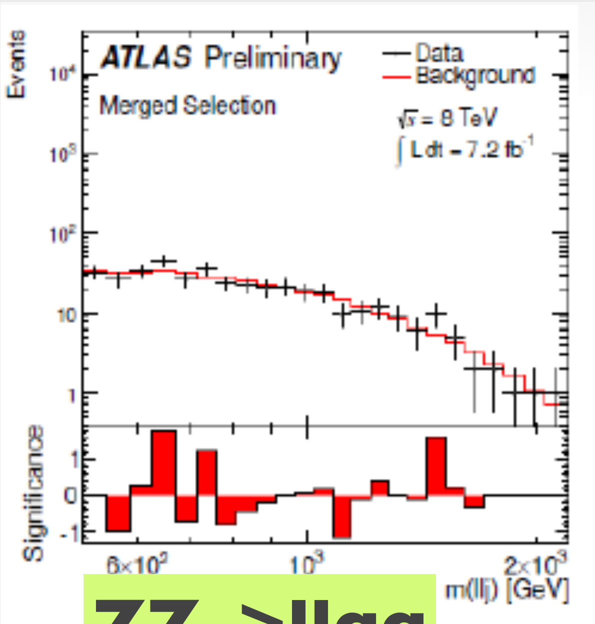
Diboson Searches



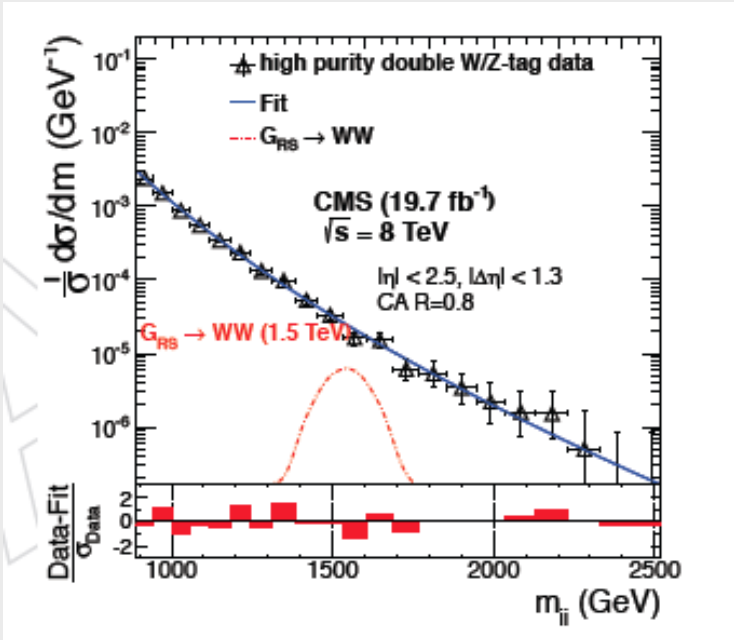
- **Search for heavy objects decaying to bosons: mixed of resolved and boosted regime (fully leptonic analysis also exist!)**
- Semi-leptonic final state: one or two leptons and 1 V-tagged jet
- All hadronic final state (VV, Vj): 1 or 2 V-tagged jet
- Data driven background estimate. Smooth background fitting + bump search in invariant M(VV)

ATLAS-CONF-2012-150

CMS-EXO-12-024
CMS-EXO-12-022
CMS-EXO-12-021



ZZ->llqq



[CMS analysis] use V-tag quality cuts to define high and low purity regions to help sensitivity at higher masses

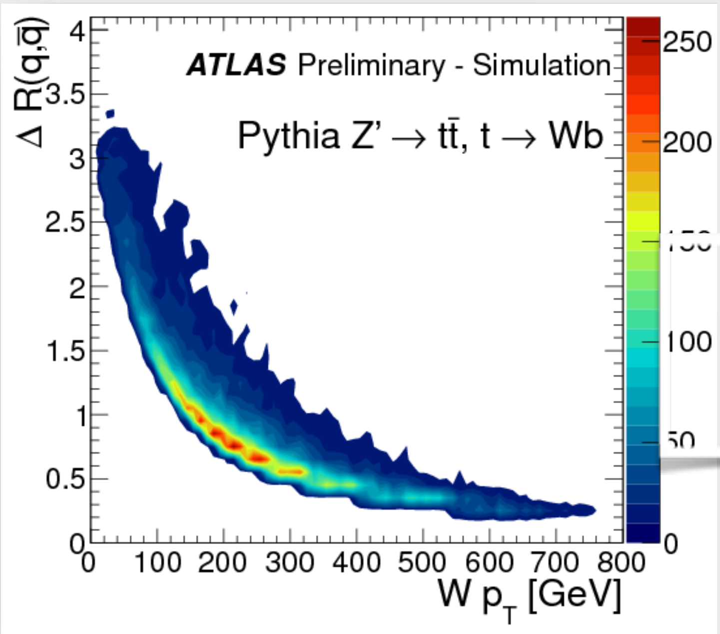
Process	Observed Mass Exclusion (TeV)
$q* \to qW$	[1.00, 3.17]
$q* \to qZ$	[1.00, 2.88]
$G_{RS} \to WW$	[1.00, 1.23]
$G_{RS} \to ZZ$	-
$W' \to WZ$	[1.00, 1.23], [1.39, 1.52], [1.57, 1.61]

X->VV (had)

ATLAS : JHEP 1309 (2013) 076

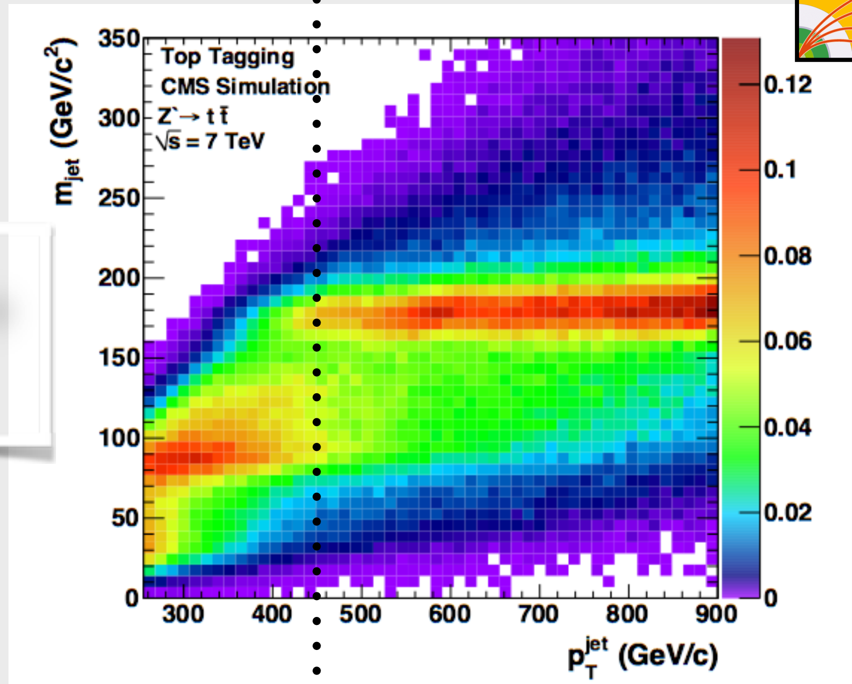
CMS-PAS-JME-13-007

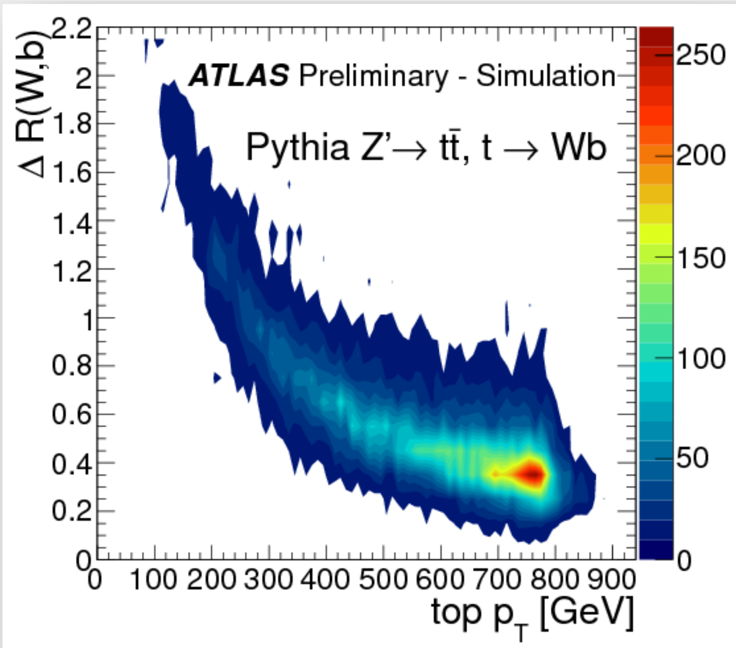
Boosted Top Kinematics



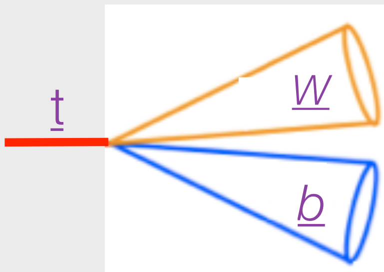
ΔR: separation in η-ϕ

ΔR ~ $2m/p_T$

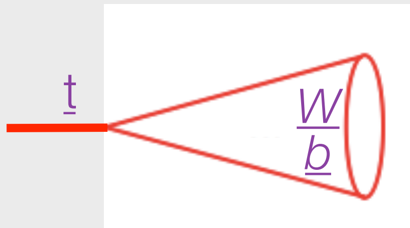


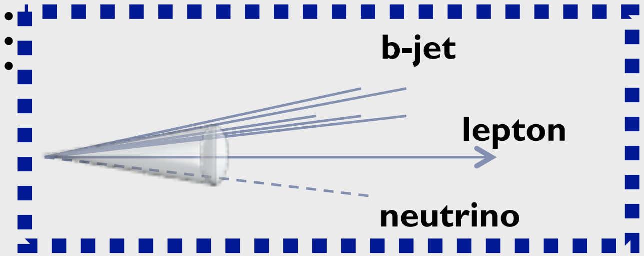


partially merged



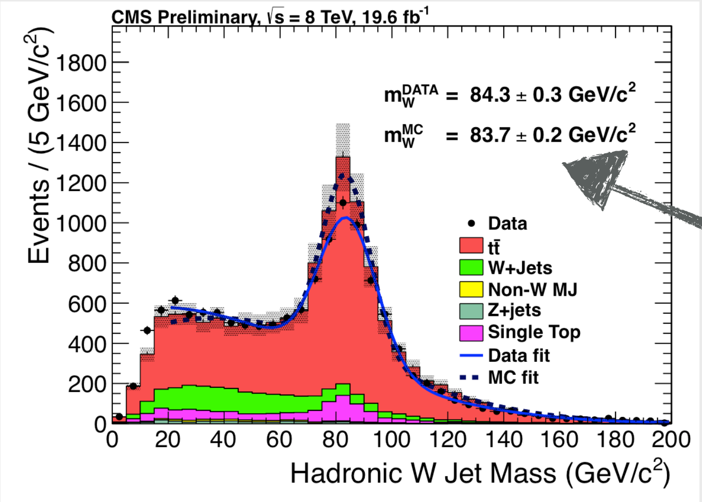
fully merged





Validation with top events in data

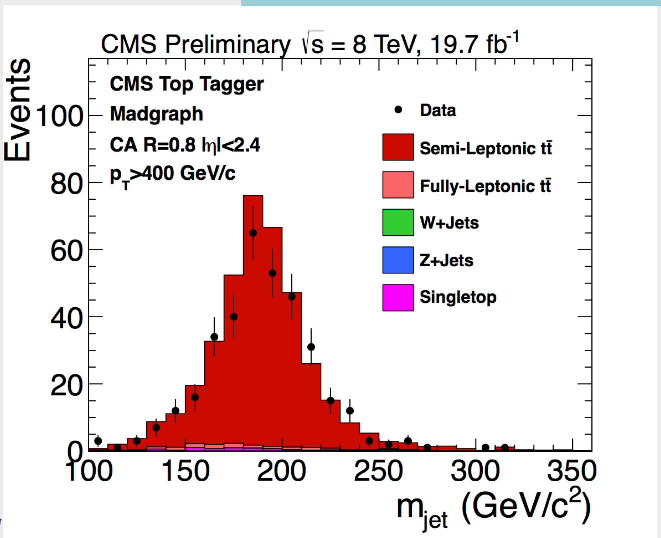
W mass in partially merged top

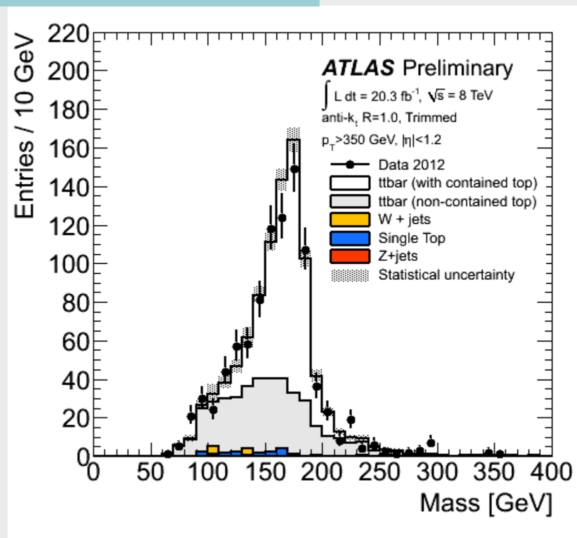


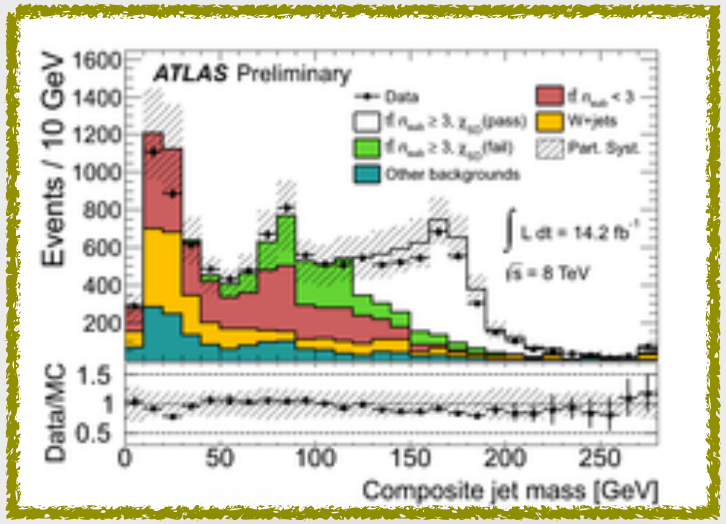
- Use tt events in data: very pure selection!
  - single lepton+MET+btag
  - -> Plus one large DR jet
- « fat » Jet mass (groomed), PU subtracted and energy corrected
- Very good agreement data/MC!

[ATLAS] New shower deconstruction algorithm: top-tagged jet used as input for SD algorithm based on parton shower model

fully merged top

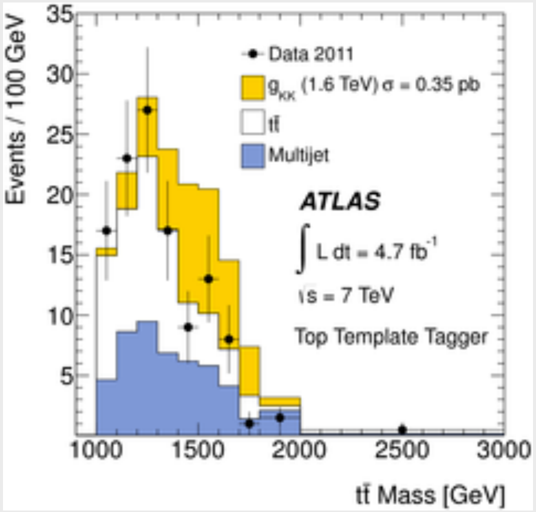


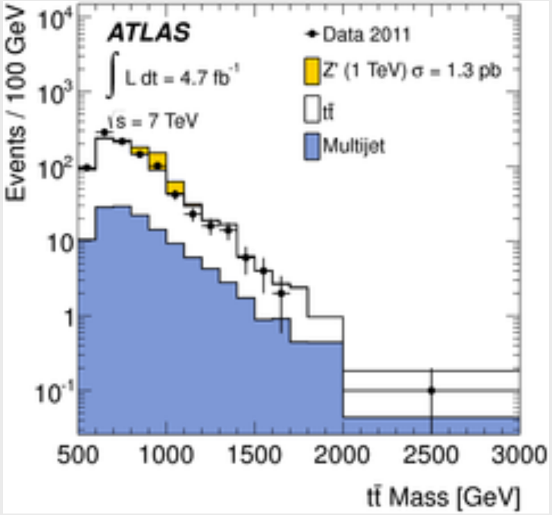


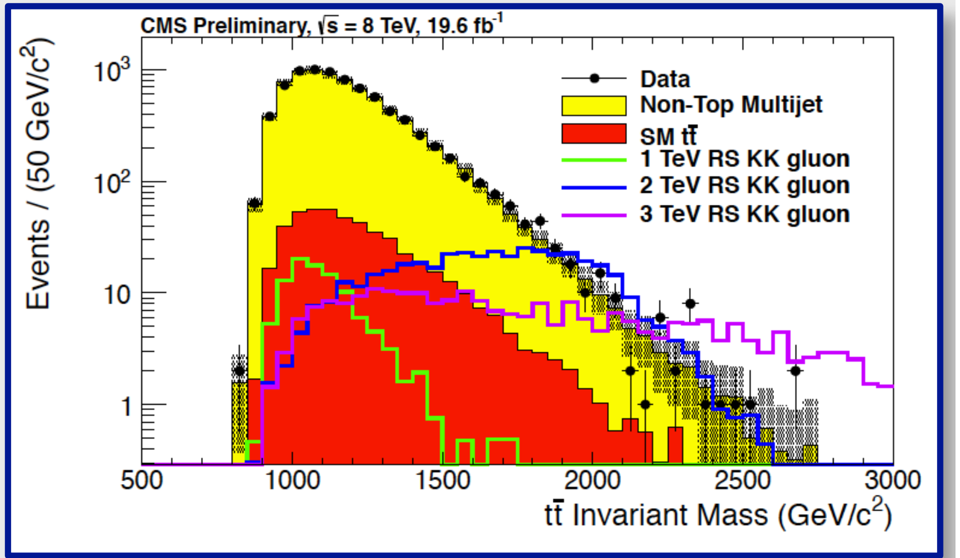


ATLAS-CONF-2014-003

Z'->tt, all hadronic







- CMS and ATLAS performed searches for heavy resonances in tt all hadronic final state: narrow and wide Z' and RS KK gluons. Mixed approach of resolved and boosted analysis.
- For the boosted case (above ~1TeV):
  - **HepTopTagger:** exploit substructure of large CA jets R=1.5. Best for pT(top)>200 GeV
  - **Top template Tagger:** exploit the energy flow of the deposits from hadronic top decay products in smaller radius R=1.0 jets. Optimized for p_T(top)>450 GeV
  - **CMS Top Tagger**: CA8 plus pruning and n-subjettiness
- CMS interprets the limit on enhanced tt production: $S=\sigma_{tt}(SM+BSM)/\sigma_{tt}(SM)<1.2$ @95%CL

CMS- PRL 111(2013) 211804

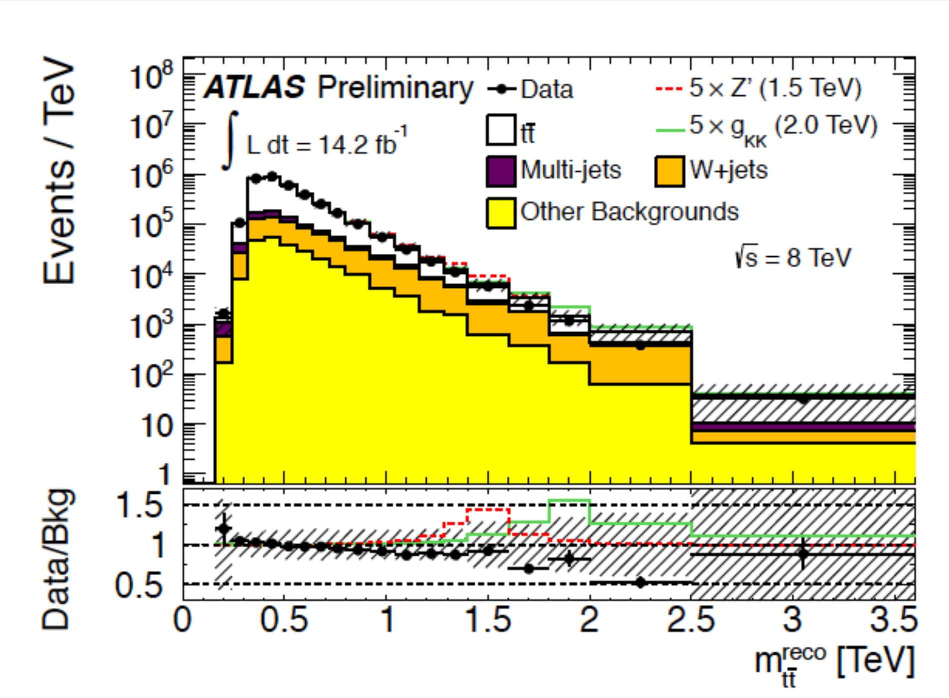
ATLAS-JHEP-01(2013) 116

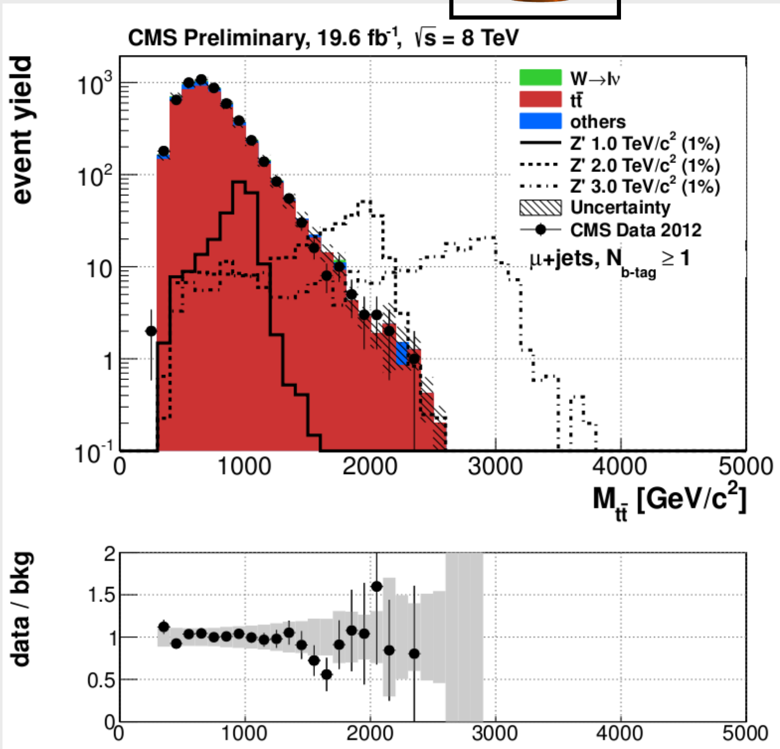
- Narrow (1%) Z' exclusion to 1.65 TeV
- Wide (10%) Z' exclusion to 2.35 TeV
- RS KK gluon exclusion to 1.8 TeV

ATLAS-CONF-2013-052

PRL 111 (2013) 211804

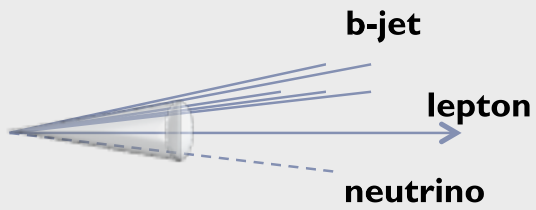
Z'->tt, l+jets





ATLAS excludes (14.3 fb^{-1}):
Narrow (1%) Z' up to 1.8 TeV
RS KK gluon up to 2.0 TeV

Mixed of resolved and boosted approach



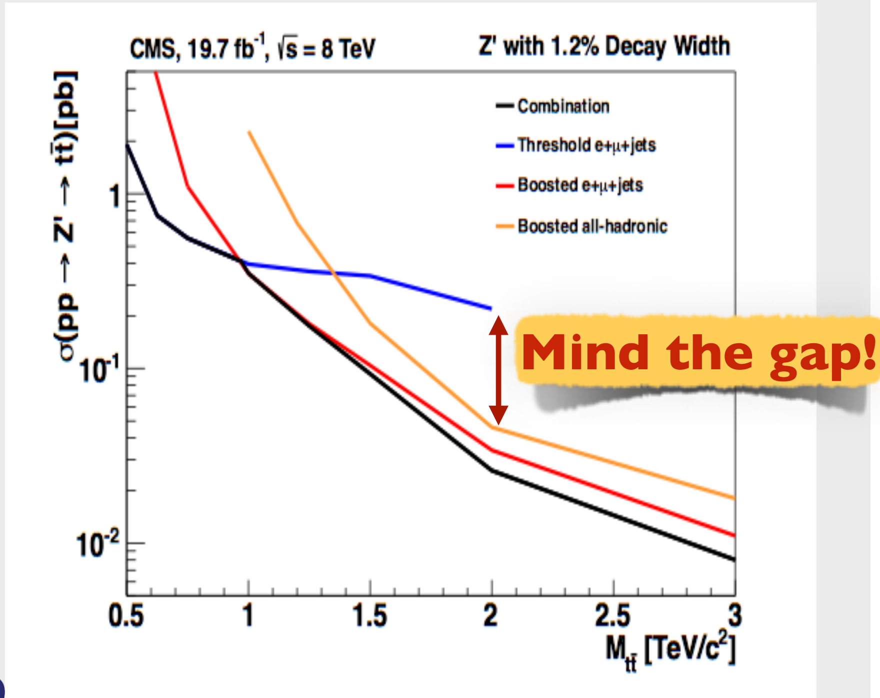
- ATLAS boosted: ≥1lepton (isolated), 1 fat jet with *mass*>100 GeV + substr., 1 small DR jet, and 1 b-tag
- CMS boosted: use *trigger with lepton+2 jets*, offline require 1 lepton and 2 jets (pt>150, 50 GeV). no isolation! split samples in number of jets and b-tags

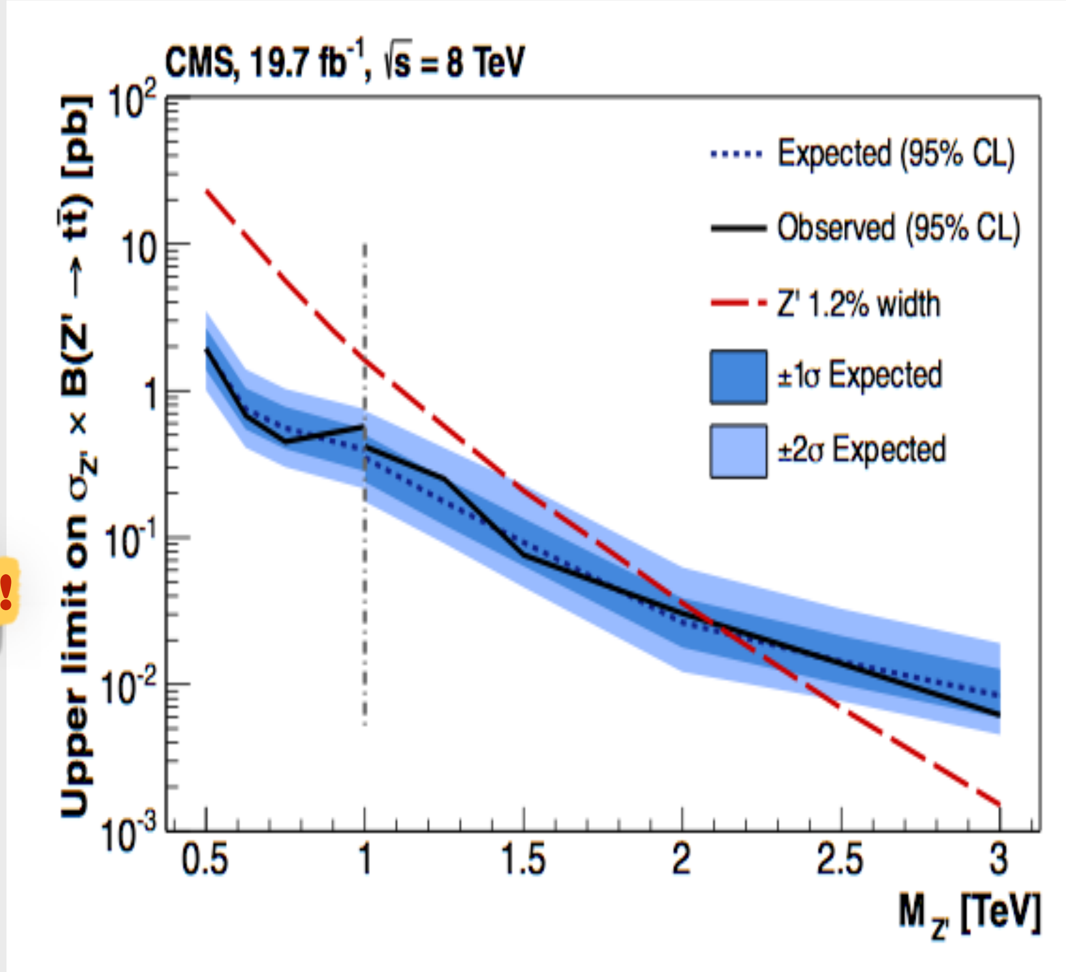
RunI combination Z'->tt (CMS)

Statistical combination of 4 channels:

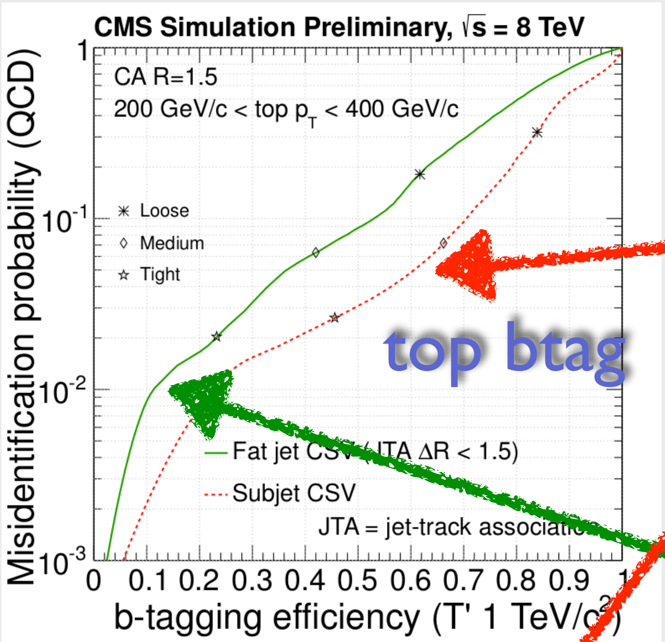
- l+j, Threshold
- l+j, Boosted 0 btag
- l+j, Boosted 1 btag
- All-hadronic

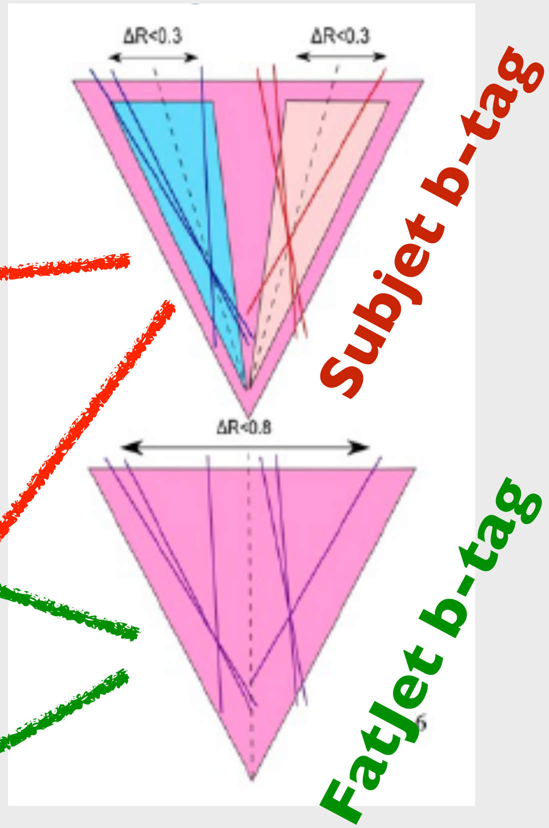
Model	Observed Limit	Expected Limit
$Z', \Gamma_{Z'}/M_{Z'} = 1.2\%$	2.1 TeV	2.1 TeV
$Z', \Gamma_{Z'}/M_{Z'} = 10\%$	2.7 TeV	2.6 TeV
RS KK gluon	2.5 TeV	2.4 TeV

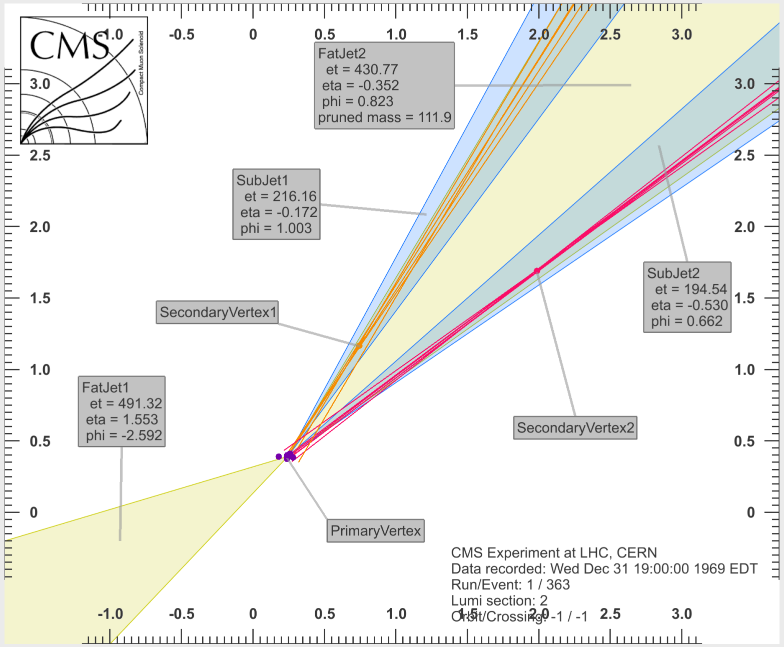


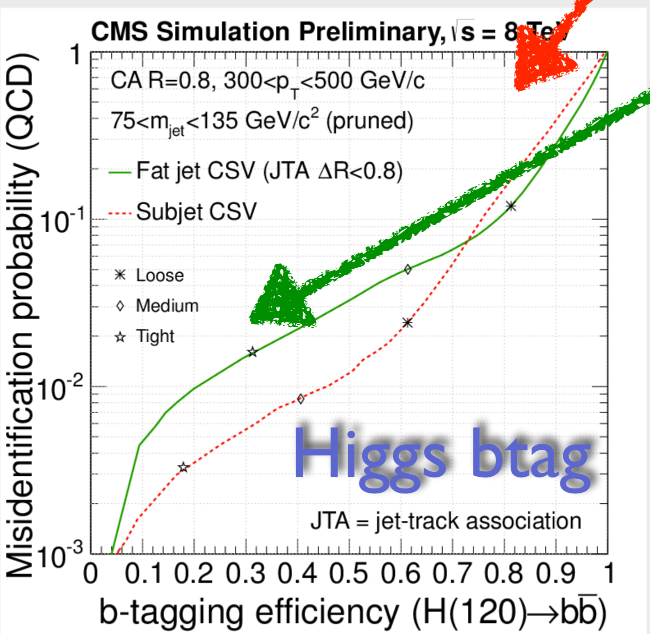


B-tagging boosted objects





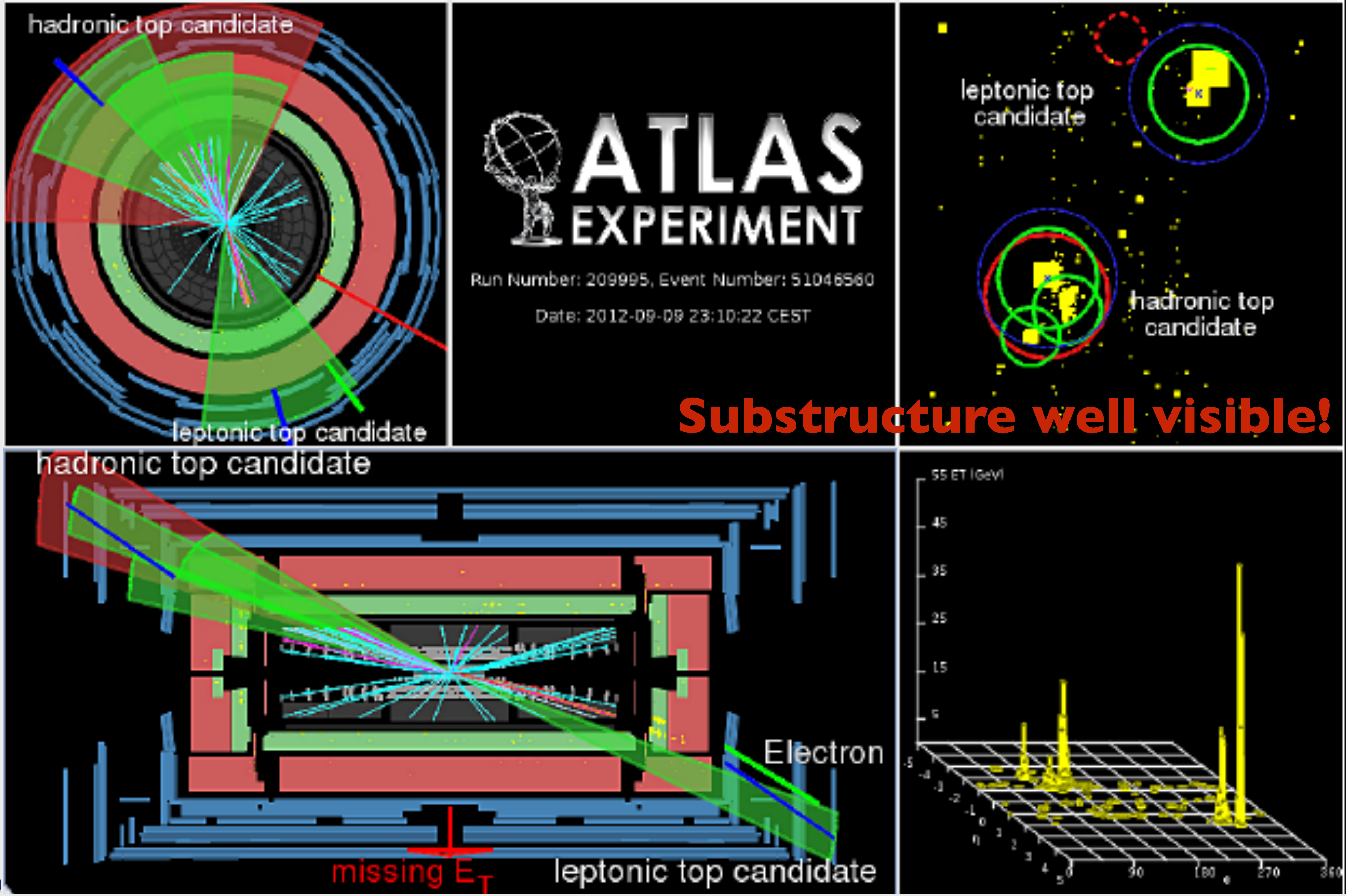




- SubJet b-tagging performs better than FatJet
- Combing with TopTagger obtain factor 10 reduction in QCD keeping 70% efficiency (compared to TopTagger only)
- At higher boosts though problems start appearing for double tag case due to large track sharing. Developments needed.

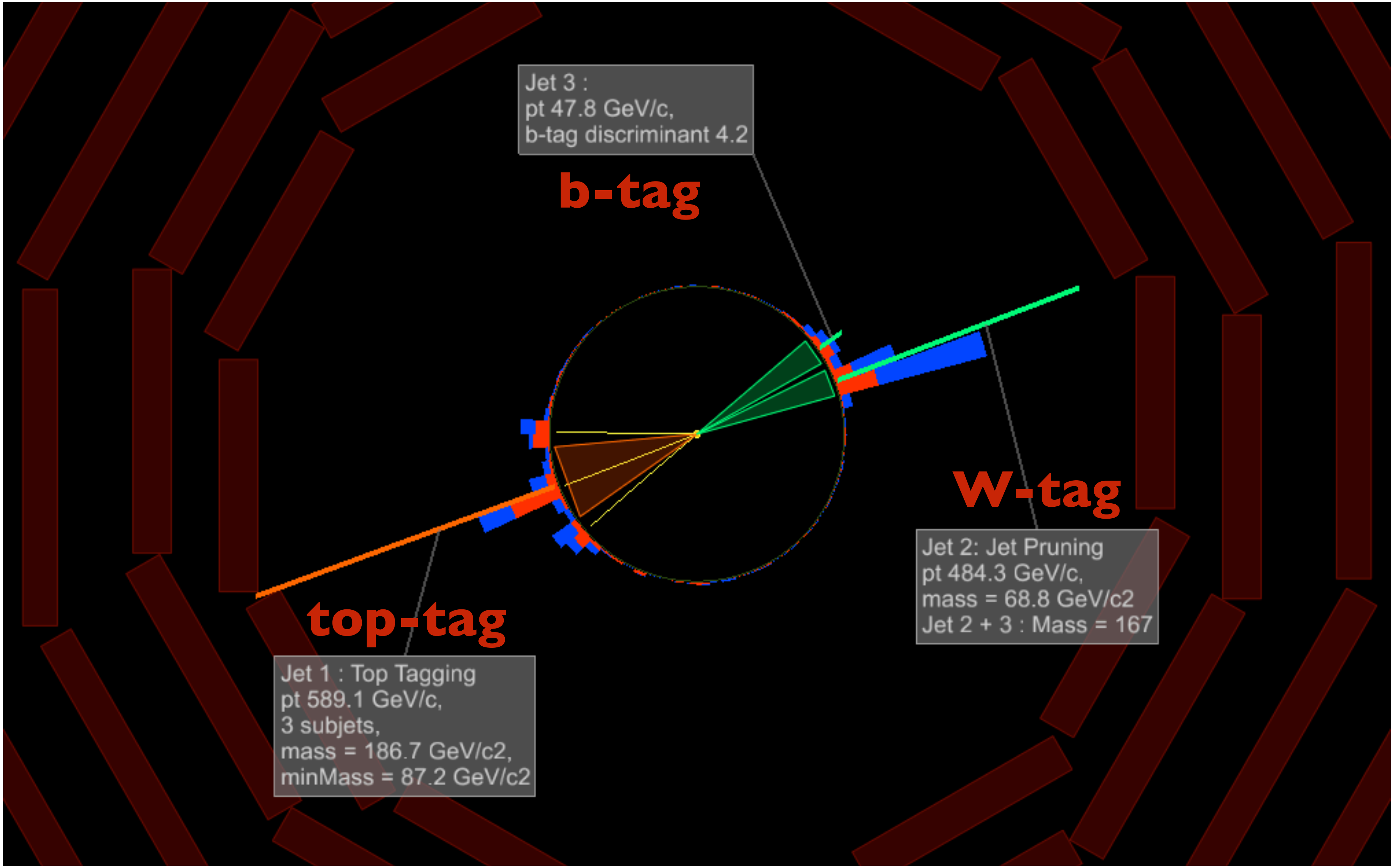
CMS-PAS-BTV-13-001

Top-tagged Event Display



Substructure well visible!

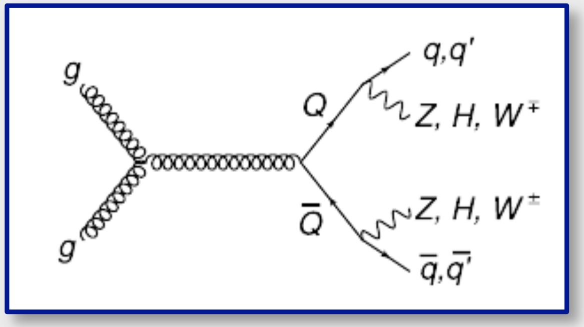
Triple-tagged top event

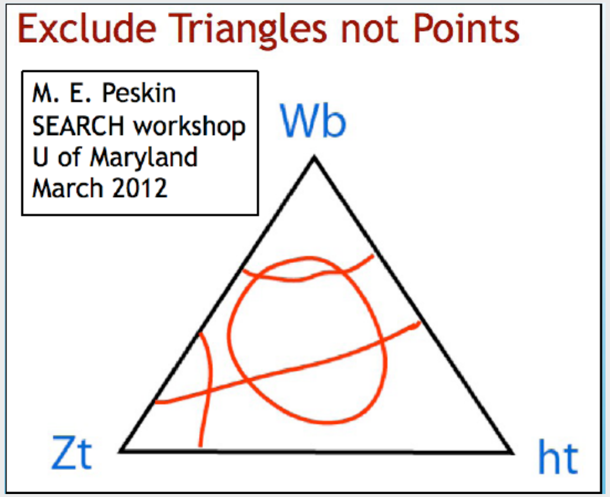


VECTOR LIKE QUARKS

- « heavier partners of the 3rd generation » appear as a very compelling extension of the SM, elaborate models still very much alive:
  - 2HDM, Little Higgs and Extra dimension predict vector-like top and bottom quark partners (not enhancing the Higgs production rate, unlike the « old » 4th generation models)
- Can have same charge as b, t (B, T) or exotic charge ($X_{5/3}$ or $Y_{-4/3}$).
- Can be isospin singlet, doublet, triplet
- Interact with 3rd-gen: mixing proportional to SM quark mass.
  - Light quark coupling sometimes enhanced.

see also talk by T. Tomei for non-boosted results





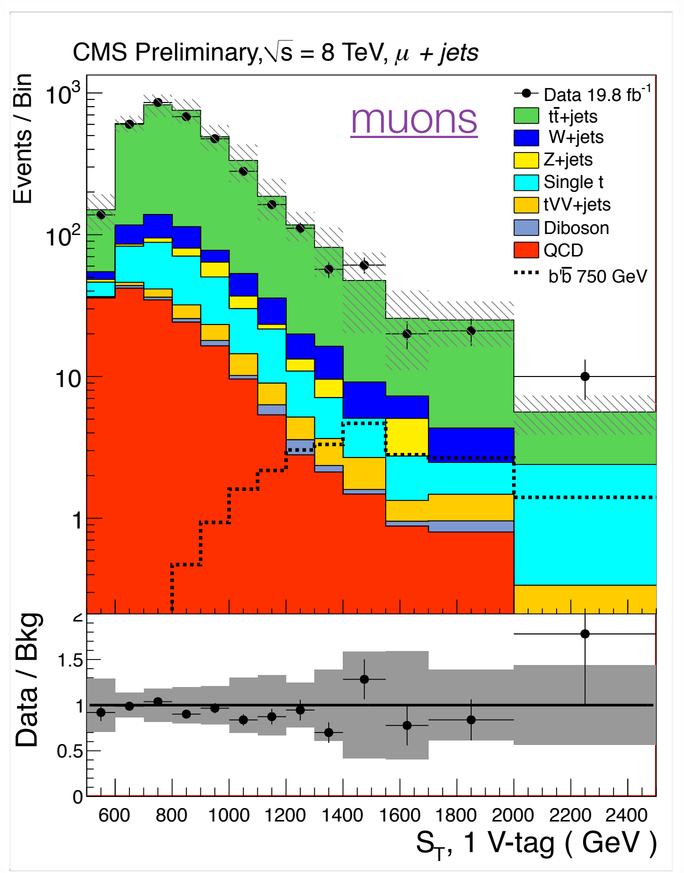
- **Many final states: B′→ tW, bZ, bH, T′→ bW, tZ, tH, $T_{5/3}$→tW**

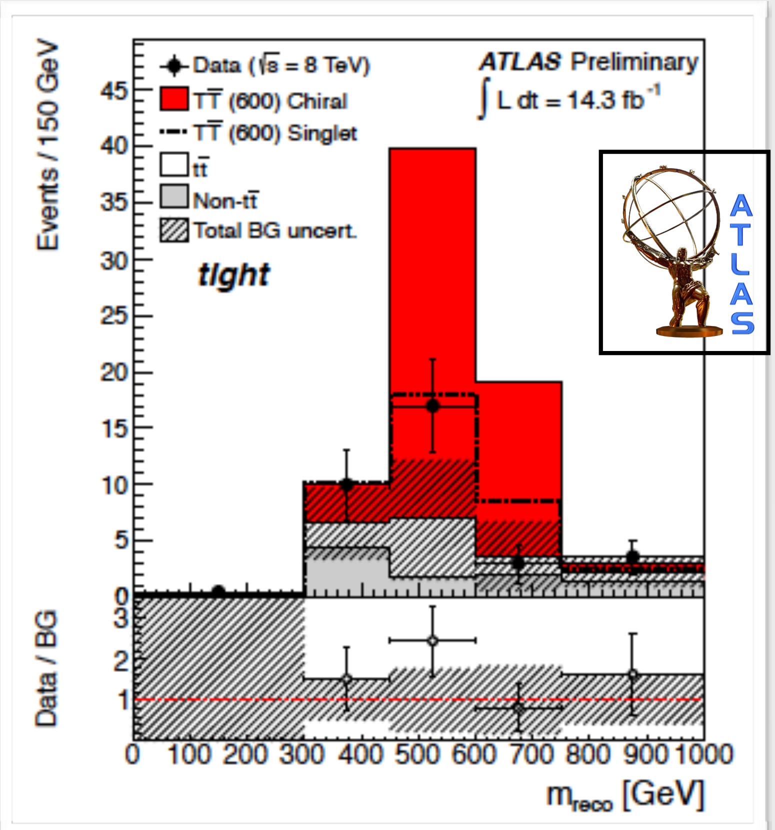
  with the constrain: BR(Wb)+BR(tZ)+BR(tH)=1
- **Leptons, b-jets, (boosted) top, (boosted) W/Z, (boosted) H,**
- **no significant Missing ET**

SEARCH FOR B' and T'

CMS: Inclusive search for B' -> tW, bZ, bH in the leptonic channel. Uses fit to total energy S_T, binned by # of boosted V-tags (=0,1,≥2). (Inclusive search for T' not shown here)

ATLAS: Inclusive search for T' -> Wb, Zt, Ht in the leptonic channel. Uses reconstructed mass of heavy object to discriminate





M(B')>732 GeV
(for 100% BR to tW)

M(T')>740 GeV
(for 100% BR to Wb)

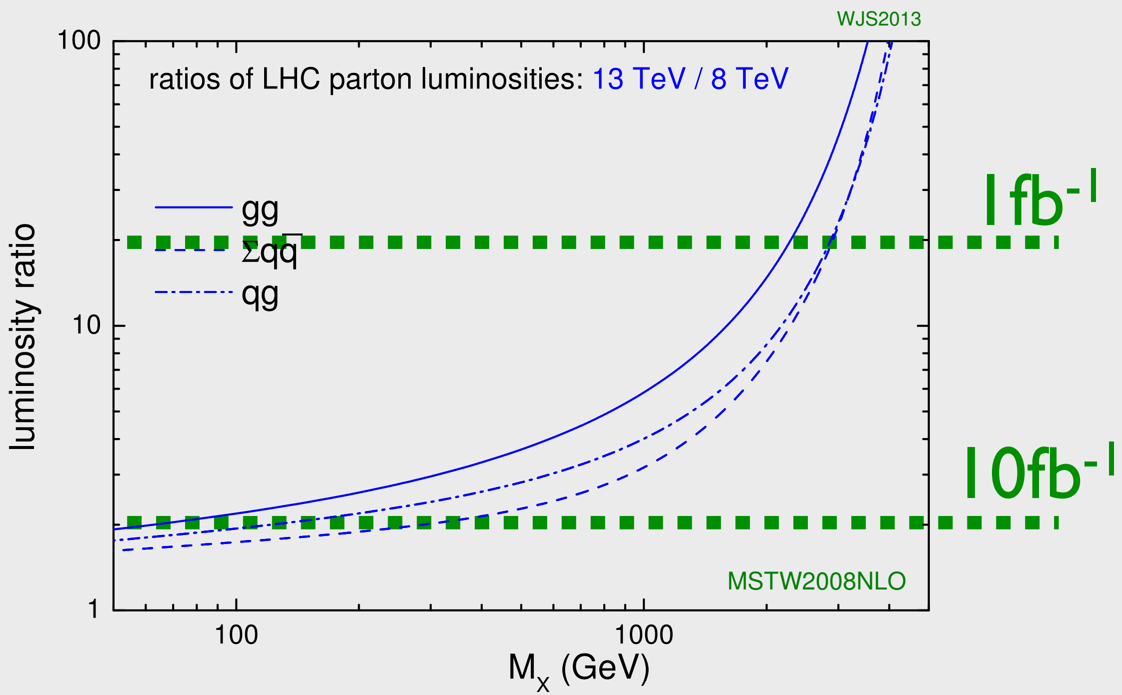
CMS hep-ex:1212:2391 submitted to PRL

CMS-PAS-B2G-12-019

CMS PLB 729 (2014)749

ATLAS-CONF-2013-060

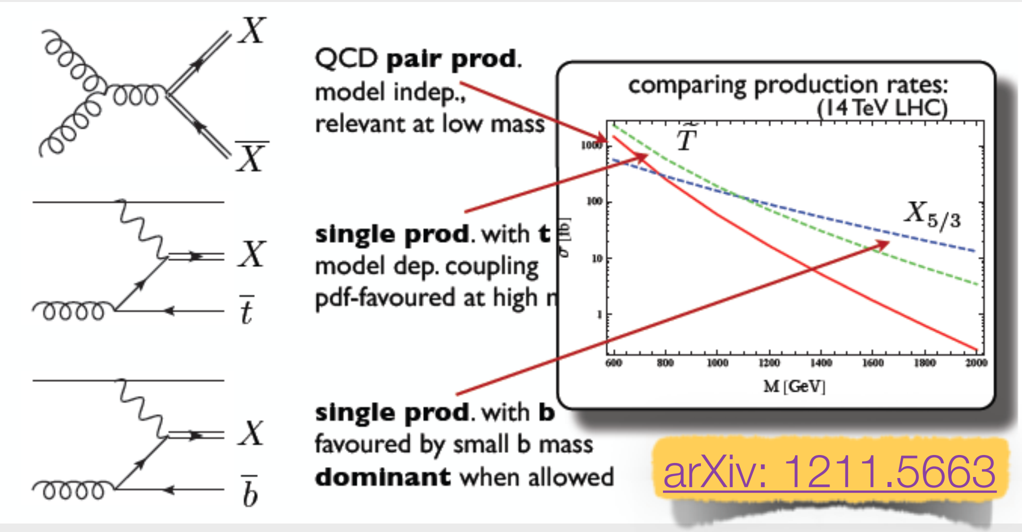
RUN II: Higher energy, Higher Pileup



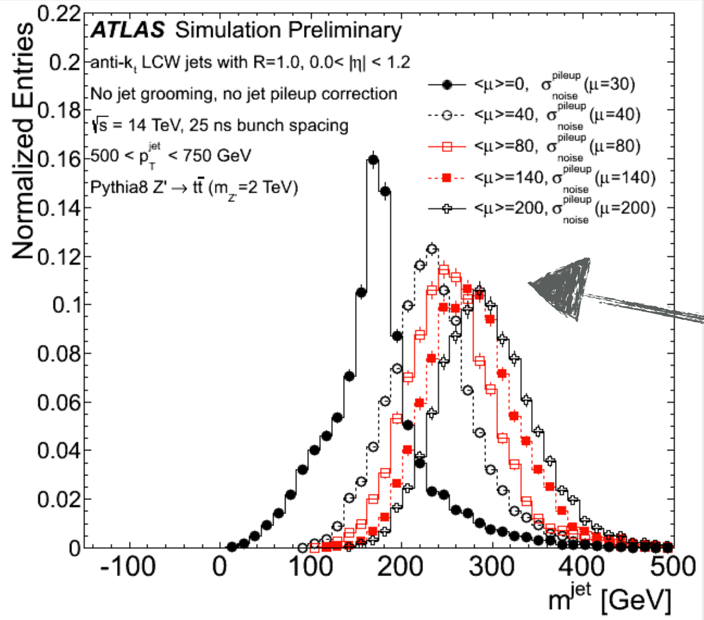
- **Higher energy+higher luminosity comes at a price: higher pileup (20—> 50 PU depending on running conditions)**
- Need new strategies to cope/profit from trigger level to analysis.
- Given Run I limits, BSM discovery geared to cover higher masses…

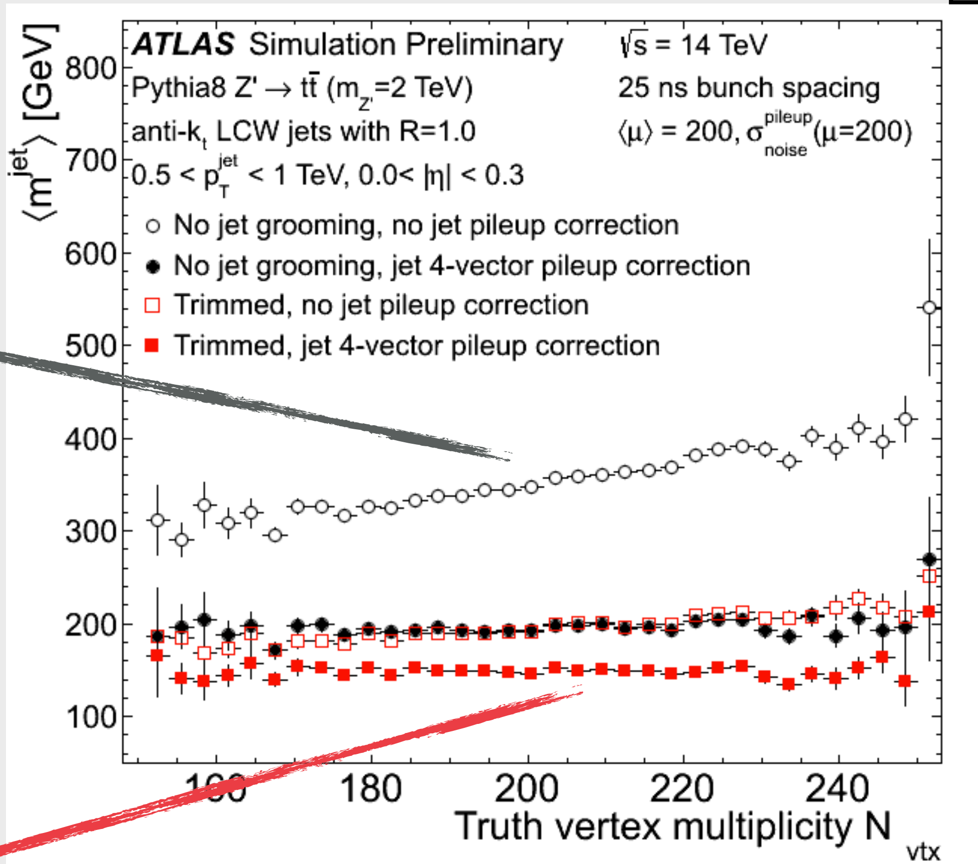
scaling of prod xs in gg fusion for $\sqrt{\hat{s}}$=2TeV, ~x15

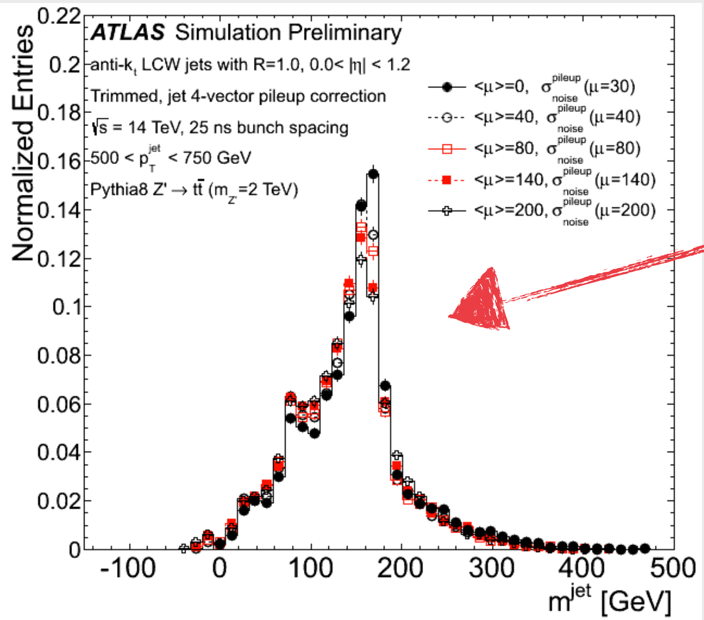
- …and new production modes becoming important! (for instance singly produced VLQ vs pair production)
- **Boosted topologies a MUST: will be widely used in trigger selection and event reconstruction**



Performance of boosted objects at high pileup







- ATLAS studied performance of large radius jets in Z'—>tt events with increasing PU conditions
- need both trimming and PU corrections to subjets to remove the <μ> dependence at very high PU.

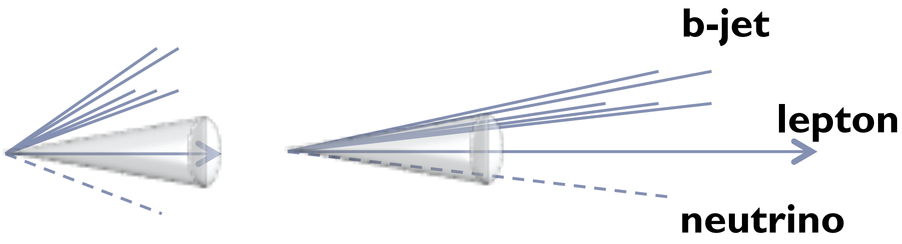
Conclusions & Prospects

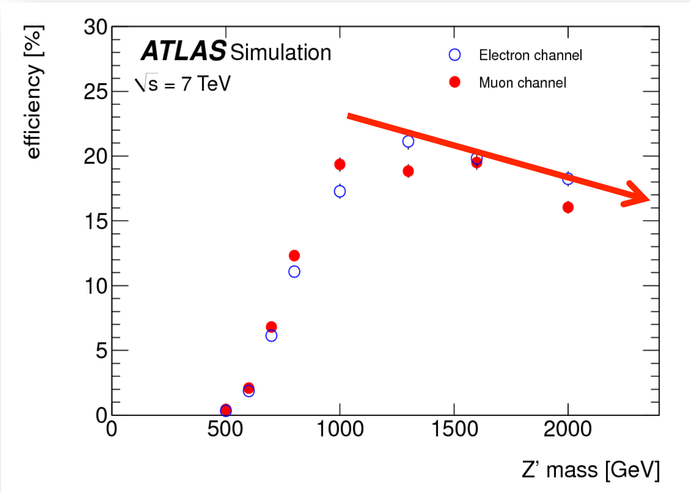
- Boosted object reconstruction is a relatively new subject that has been now widely recognized as crucial for ATLAS and CMS:
    - **it has allowed to increased significantly the analysis sensitivity for heavy object searches (Z ‘, VLQ…)**
    - **but for SM measurement as well for W/Z/H/top**
- **This improved sensitivity will allow to cover new areas of the phase space for new signal production modes enhanced by the higher energy available in Run II**
    - for instance single production of heavy objects that can be recast in many models
- As a byproduct this new methods provide also:
    - very stable performance against pileup
    - clever new ways to reduce the SM backgrounds
    - extremely useful for studying tail of SM processes
- **Let’s not forget that the application of these latest techniques on the full Run I datasets might still bring surprises to the legacy papers currently in preparation by CMS and ATLAS!**

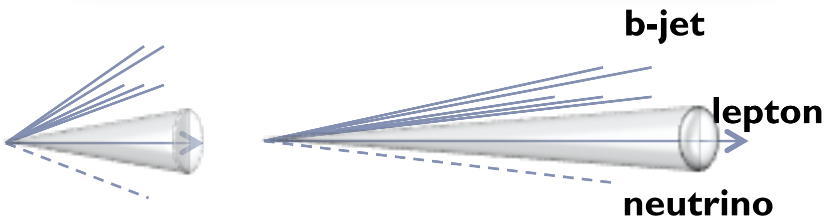
BACKUP SLIDES

Lepton Isolation

- lepton and b-jet from boosted top highly collimated: lose isolation efficiency.
- But even boosted, leptons from tops have larger separation than those from light quark jets.
- Loss in efficiency can be recovered in part: variable p_T-dependent cone size, "mini isolation", ...

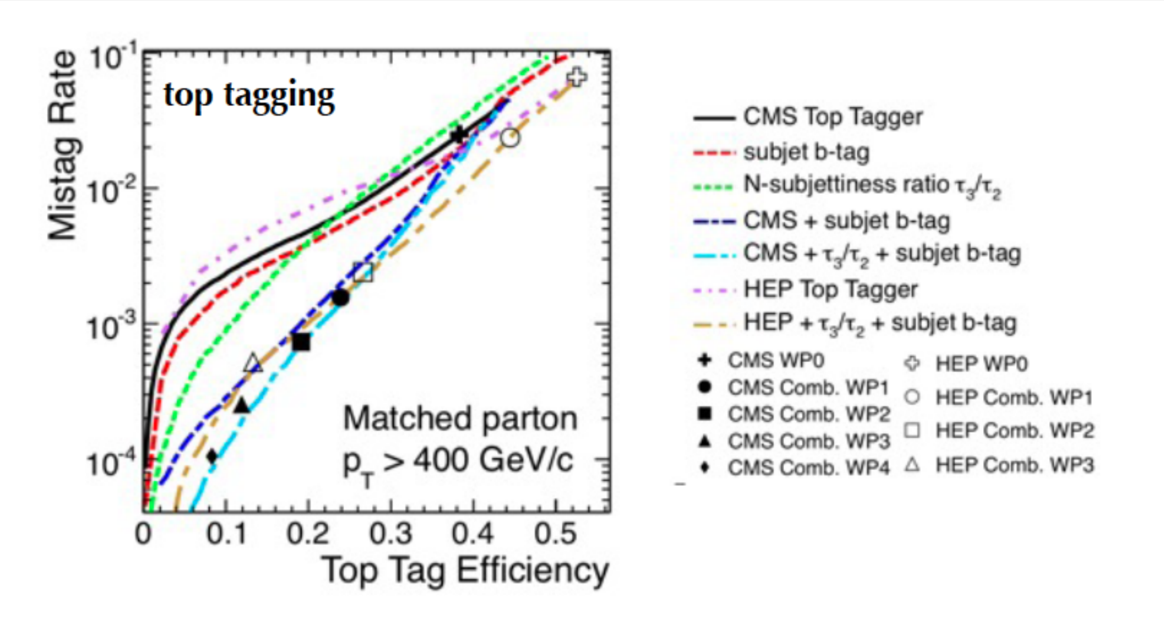


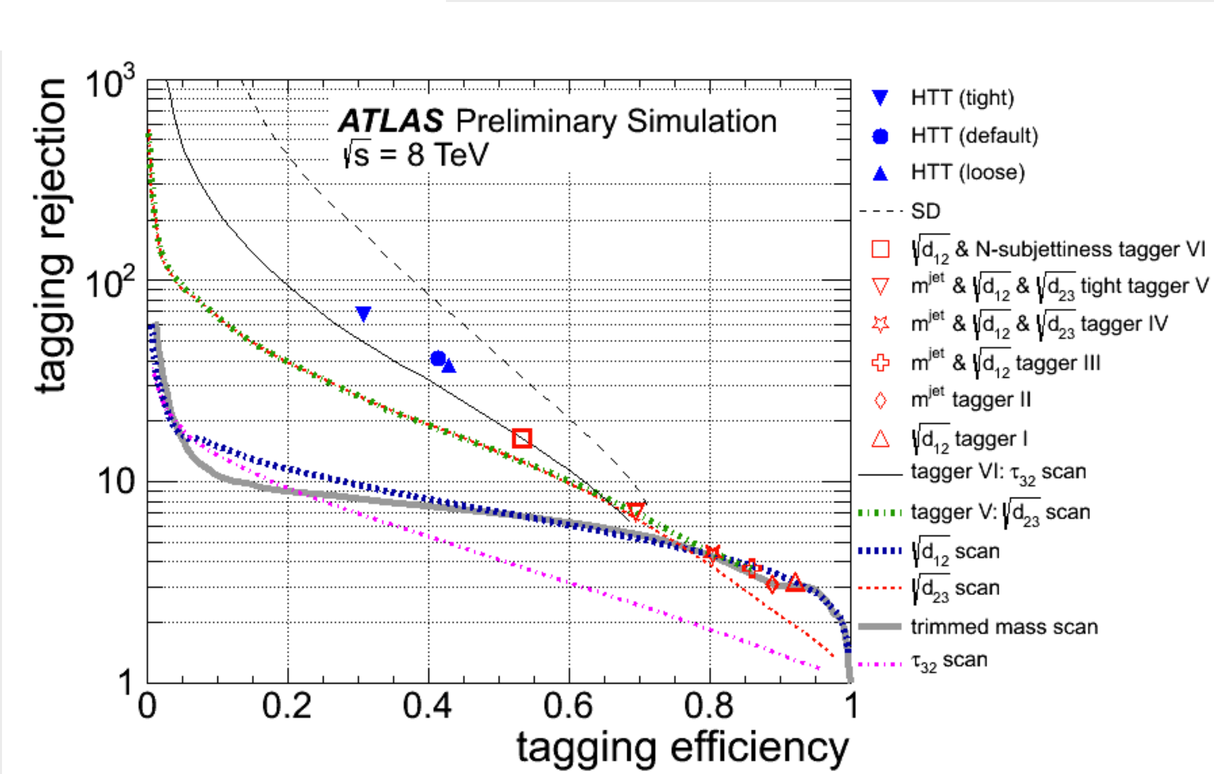




JHEP 1103:059 (2011)
Rehermann, Tweedie

top tagging ROC curves



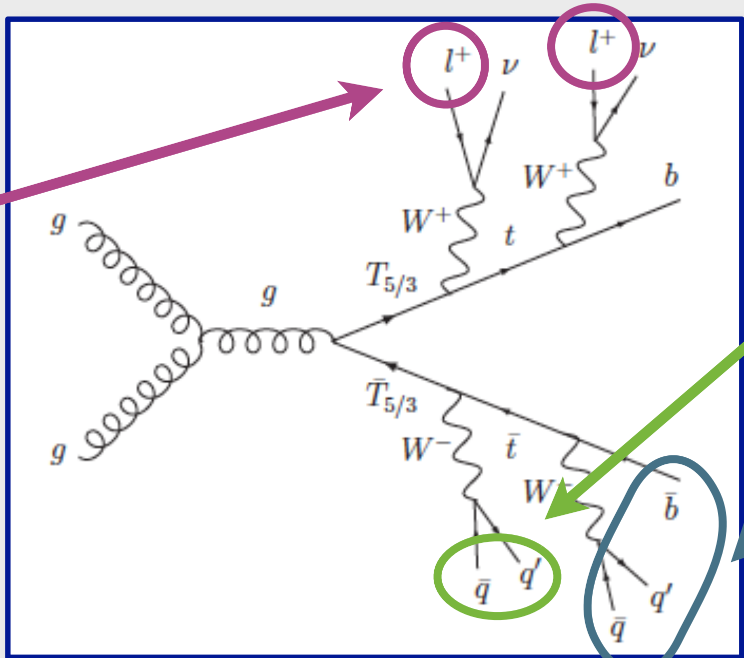


Pair production of $T_{5/3}$

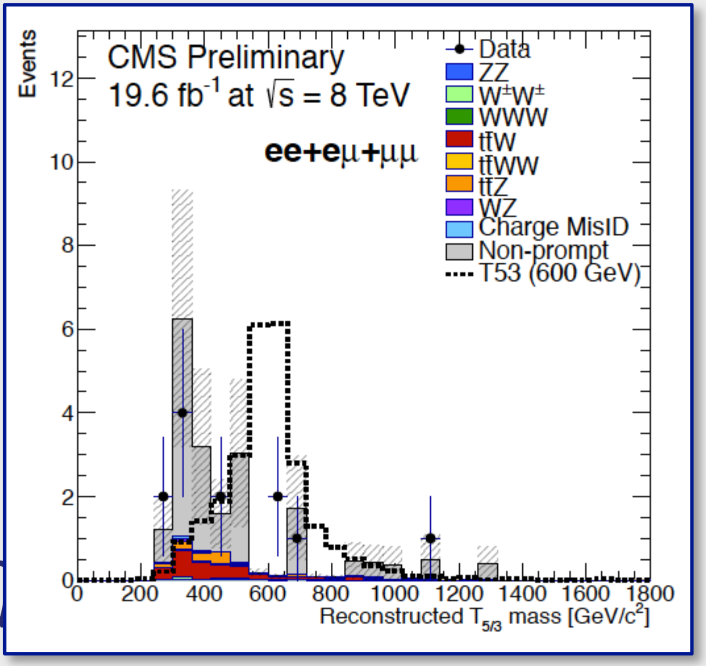
hep-ex:1212:2391
submitted to PRL

same-sign dileptons

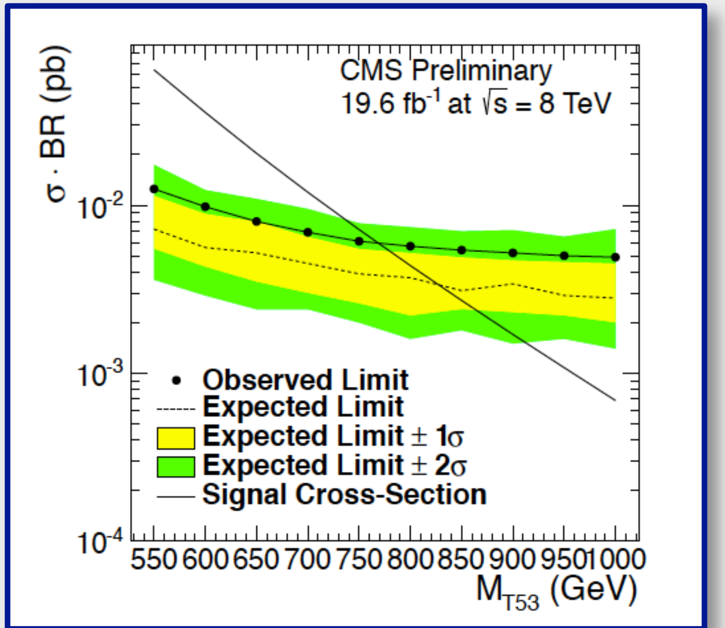
maybe skip?



- boosted W tag: 2 subjets, M_j[60,130]
- CMS top tag
- reach 13% lower xs using boosted reco

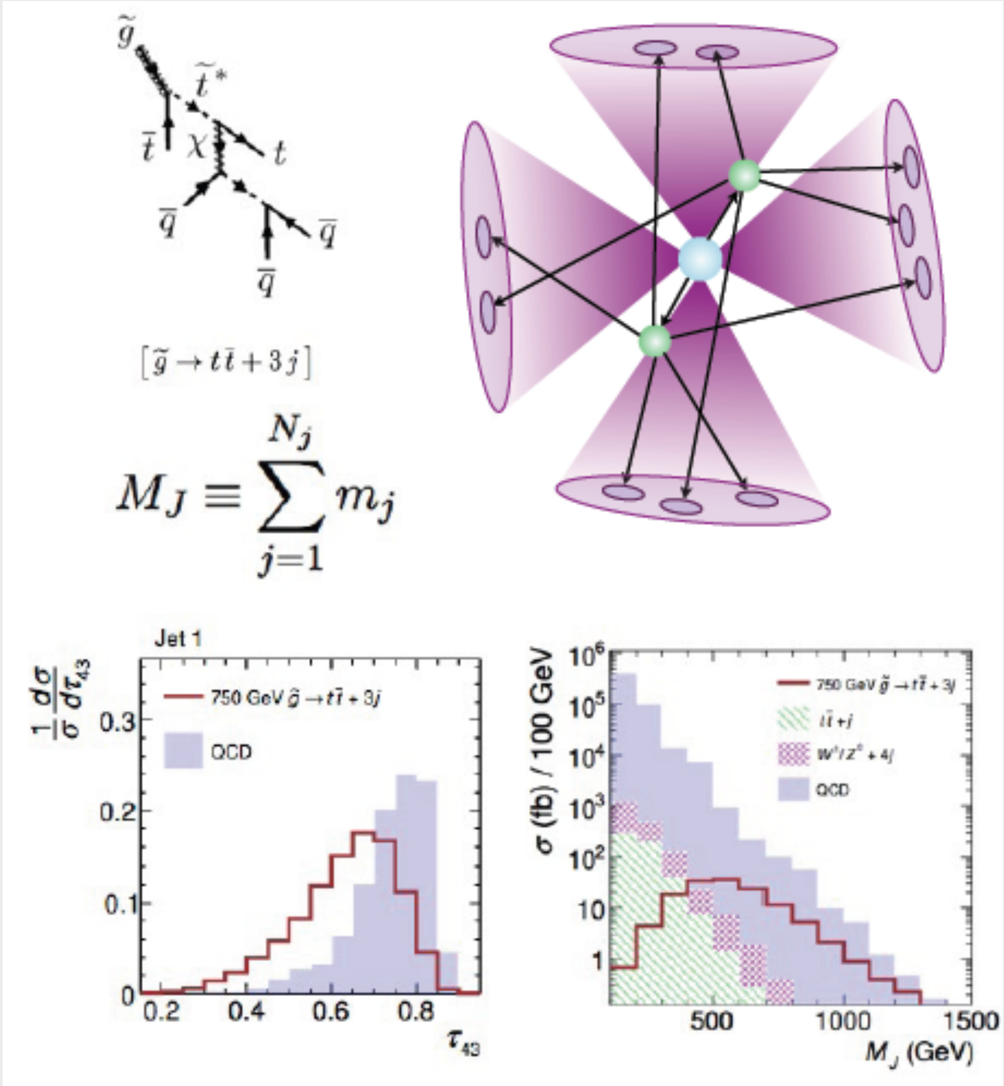


$M_{T(5/3)}$ > 770 GeV



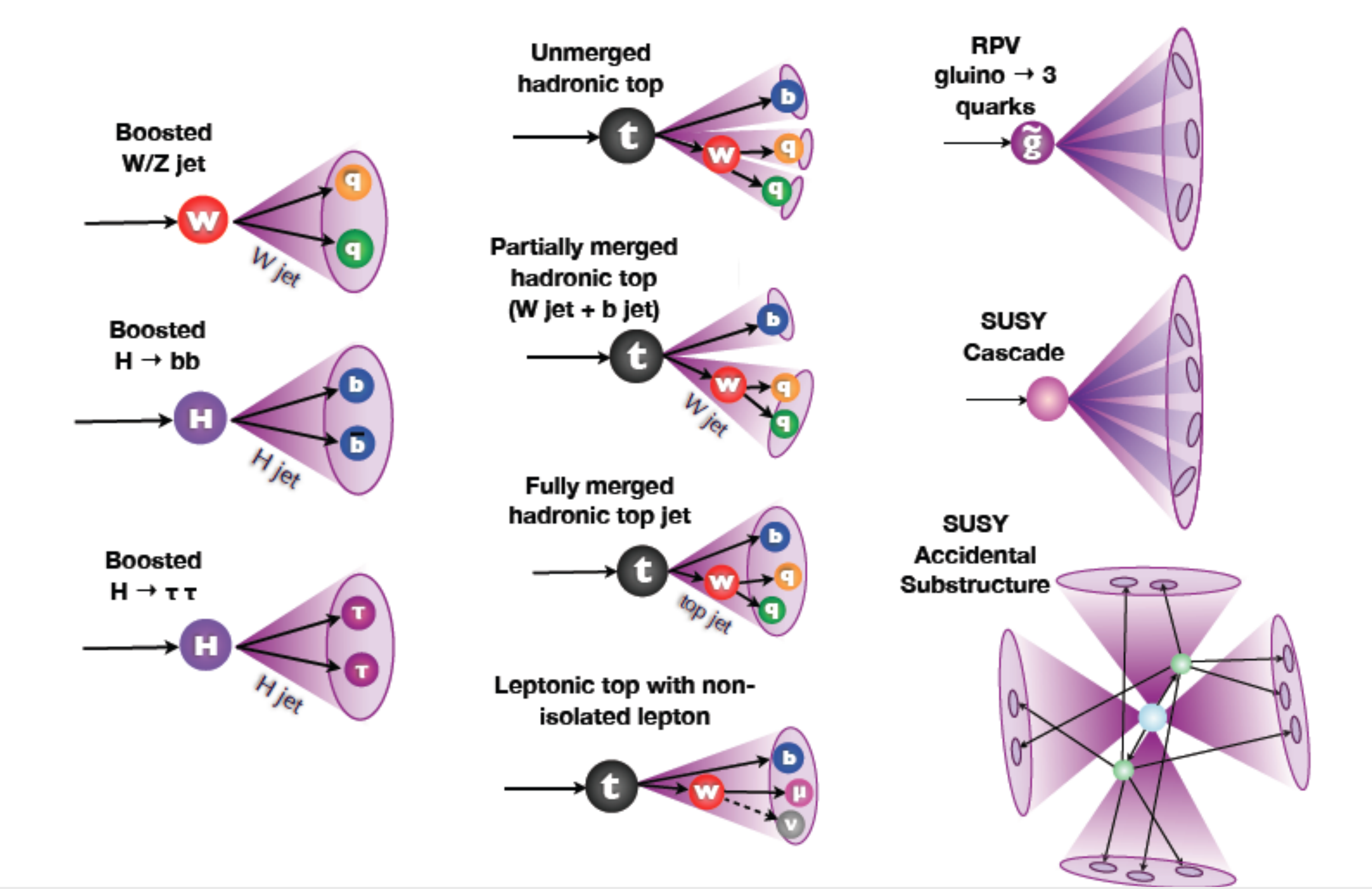
SuSY accidental substructure

- Another interesting example concerns the possibility of fat jets with « accidental substructure »
- Here there is no real « boost » but only a large multiplicity of particles for instance from Susy cascade decays
- These particles spread out in a way that allows them to be reconstructed together with a single large radius jet.
- the same substructure tools can be used to reduce the background for instance from QCD processes and other SM



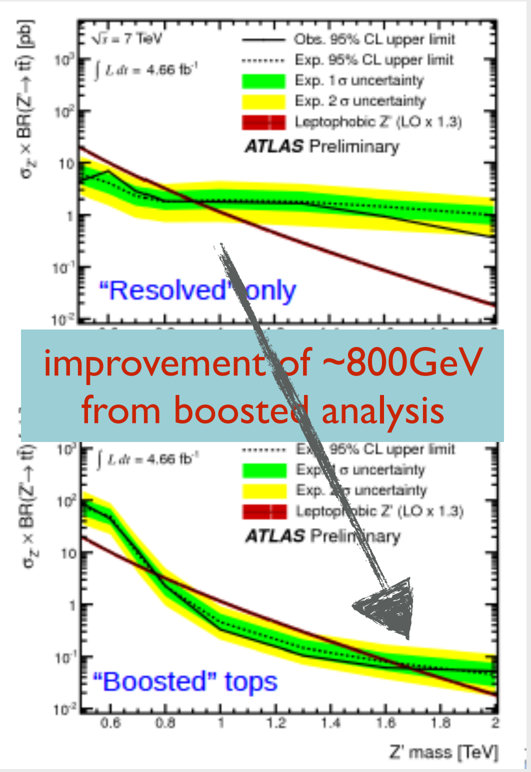
arXiv:1212.1456

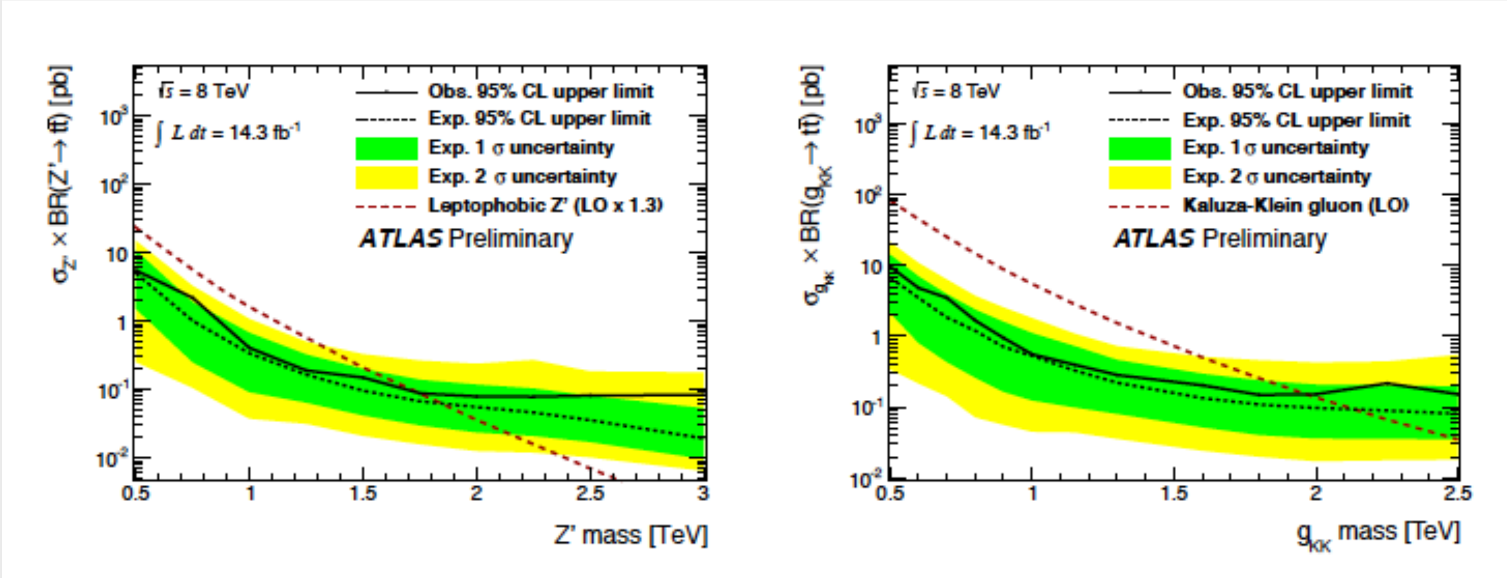
Boosted object menu



Z'->tt, lepton+jets

7 TeV data





Combined 95% limit @8TeV

0.5 TeV < m(Z')<1.8 TeV   0.5 TeV< m(g_{KK}) < 2.0 TeV

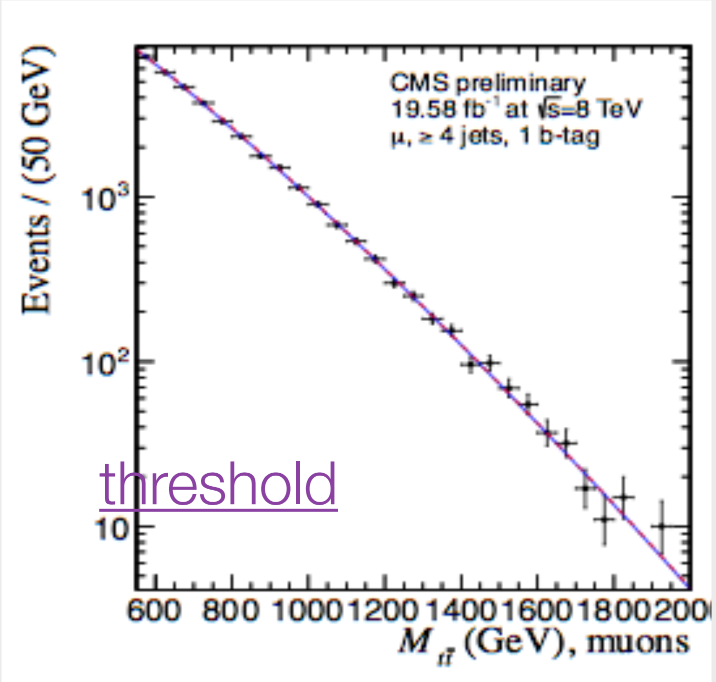
- Top jet reco: anti-Kt R=1.0, trimmed, p_T>350 GeV, M(jet)>100GeV. Splitting scale √(d12)>40 GeV
- Boosted selection require « fat » jet to be far from lepton and from other small R jet. Presence of a b-tag in small R jet is required.
- Data driven background estimate

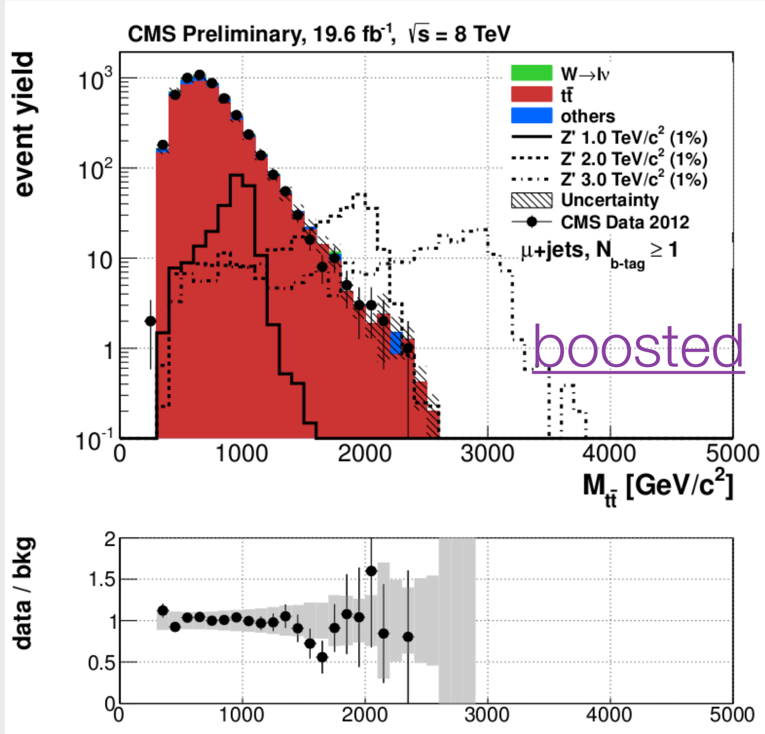
ATLAS-CONF-2013-052

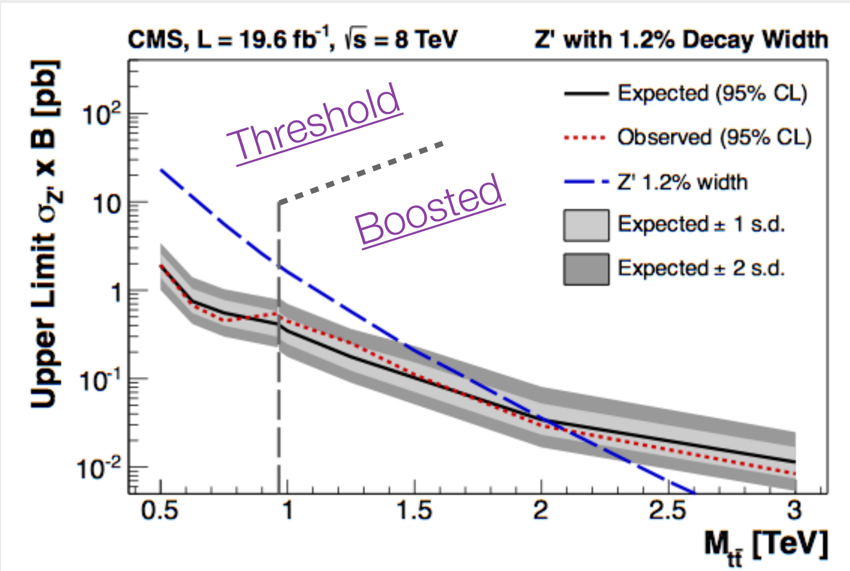
Phys.Rev. D88 (2013) 012004

JHEP 1301 (2013) 116

Z'->tt, lepton+jets



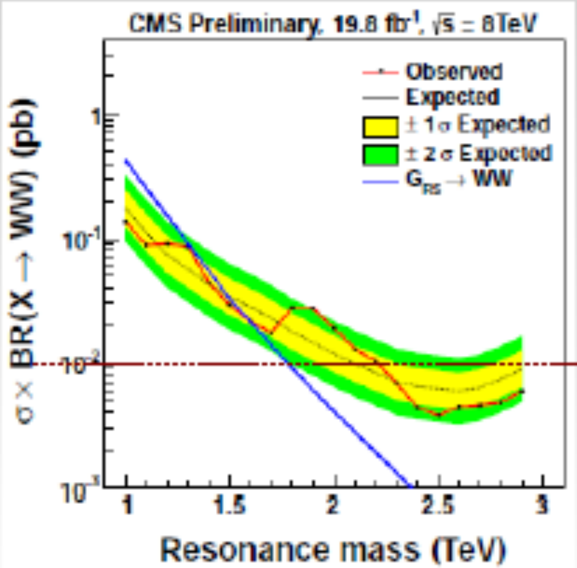


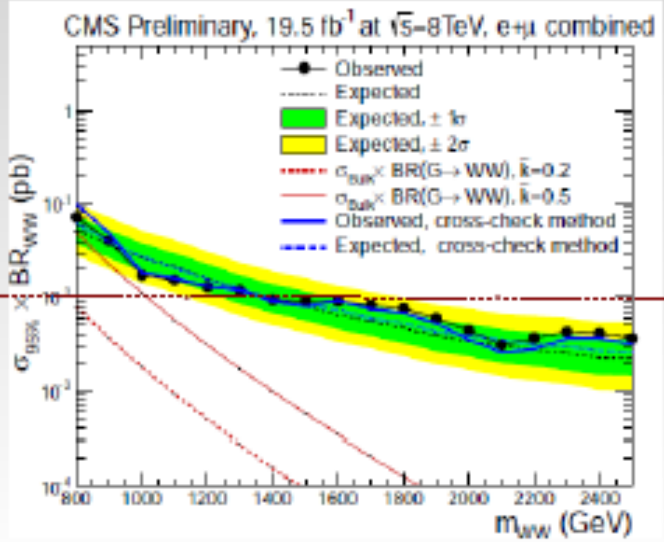


CMS excludes:

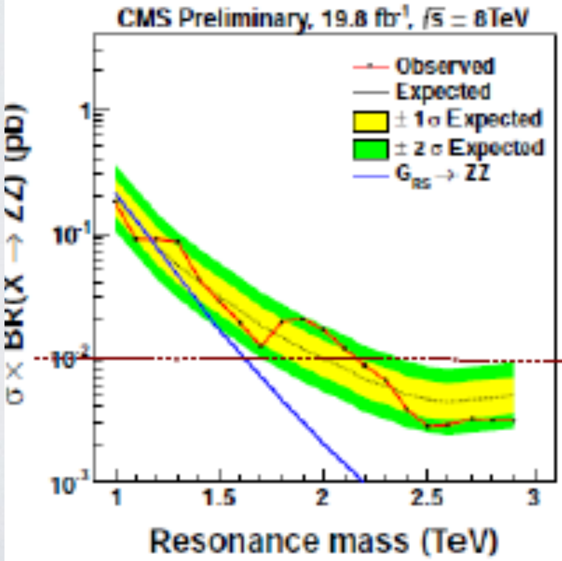
- Narrow (1%) Z' up to 2.1 TeV
- Wide(10%) Z' up to 2.68 TeV
- RS KK gluon up to 2.5 TeV

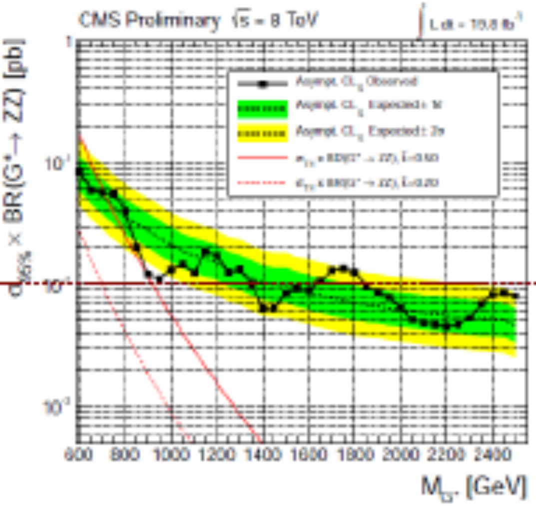
CMS-PAS-B2G-12-006

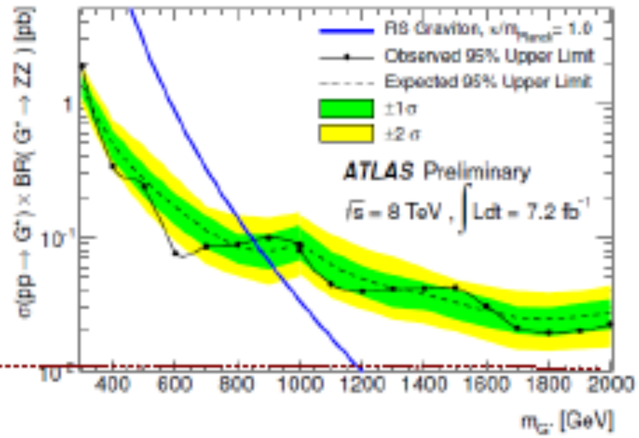




WW hypothesis







ZZ hypothesis

W jet tagging details

- **Mass drop, μ:** Two subjets are obtained by undoing the last clustering iteration of the pruned jet clustering to obtain the variables used in Ref. [40]. The ratio of masses of the highest mass subjet (m_1) and the total pruned jet is defined as the *mass drop* $\mu = \frac{m_1}{m_{\rm jet}}$.
- **N-subjettiness, τ_N:** N-subjettiness was introduced in Ref. [41] and is a generalized jet shape observable. For N candidate subjets of a given jet, we can define the N-subjettiness observables as:

$$\tau_N = \frac{1}{d_0}\sum_k p_{T,k}\min\{\Delta R_{1,k}, \Delta R_{2,k}, \cdots, \Delta R_{N,k}\} \qquad (1)$$

where k runs over all constituent particles. The normalization factor is $d_0 = \sum_k p_{T,k} R_0$ and R_0 is the original jet radius. The τ_N observable has a small value if the jet is consistent with having N subjets. Thus, for boosted W identification, i.e. for discrimination of W-jets with 2 subjets and QCD jets consistent with 1 subjet, the ratio τ_2/τ_1 is of particular interest and tends to smaller values for signal W-jets. The subjet axes can be optimized to minimize the N-subjettiness value. As a default definition for the axes, we use a one step optimization of the exclusive k_T axes.

- **Qjet volatility, $\Gamma_{\rm Qjet}$:** Qjets was introduced in Ref. [42] as a statistical interpretation of jet trees. A typical jet tree is defined by its cluster sequence; however, the jet can be reinterpreted as a distribution of trees. The process for deriving a distribution of trees is defined by: (1) at every state of clustering, assign a weight to each constituent pair, w_{ij} and (2) generating a random number to choose the $2 \to 1$ clustering from the available pairs. The default weight is defined as:

$$w_{ij} = \exp\{-\alpha\frac{d_{ij} - d^{\rm min}}{d^{\rm min}}\} \qquad (2)$$

where d_{ij} is the distance between the ij pair. As an example, for the k_T algorithm, $d_{kT} = \min\{p_{T,i}^2, p_{T,j}^2\}\Delta R_{ij}^2$ and for the CA algorithm, $d_{\rm CA} = \Delta R_{ij}^2$, where the CA algorithm is the default. From this distribution of $N_{\rm trees}$, which we choose to be 50, one can extract the Qjet volatility which is defined as the RMS of the distribution over the average jet mass, $\Gamma_{\rm Qjet} = {\rm RMS}/\langle m\rangle$. In order to improve the speed of the algorithm while not degrading the performance greatly, before running the Qjet clustering, we pre-cluster the jet constituents down to 35.

- **Generalized energy correlation functions, C_2^β:** Observables with n-point correlation functions have been proposed in Ref. [43]. The 3-point correlation function is particularly useful for W-tagging:

$$C_2^\beta = \frac{\sum_{i,j,k} p_{Ti}p_{Tj}p_{Tk}(R_{ij}R_{ik}R_{jk})^\beta \sum_i p_{Ti}}{(\sum_{i,j} p_{Ti}p_{Tj}(R_{ij})^\beta)^2} \qquad (3)$$

- **Jet charge, Q^κ,** The jet charge [44] is a measure of the electric charge of the particle originating the jet. It is defined as:

$$Q^\kappa = \frac{\sum_i q_i (p_T^i)^\kappa}{(p_T^{jet})^\kappa} \qquad (4)$$

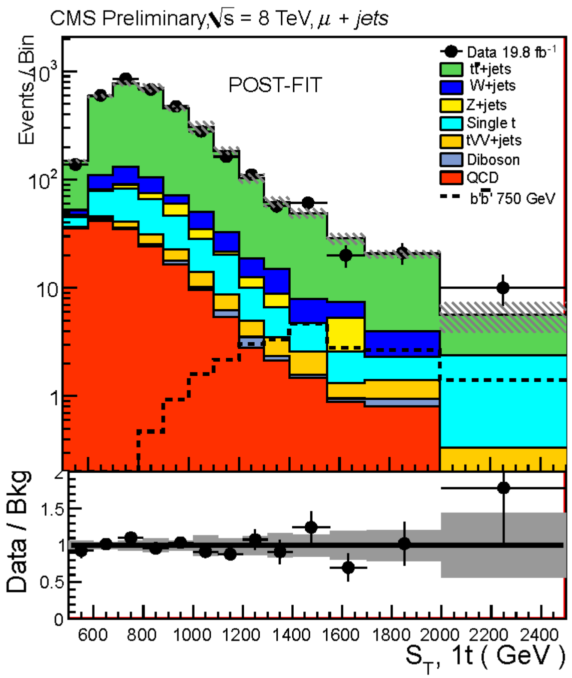
where i runs over all particles in a jet. It can be used to provide additional discrimination between quark jets and gluon jets or also to distinguish between, for example, a charged W′ and a Z′ new physics signal.

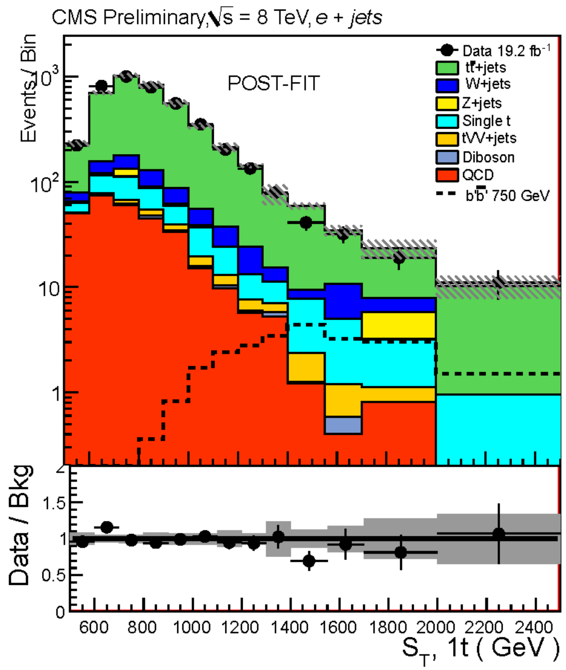
$$Q^\kappa = \frac{\sum_i q_i (p_T^i)^\kappa}{(p_T^{jet})^\kappa}$$

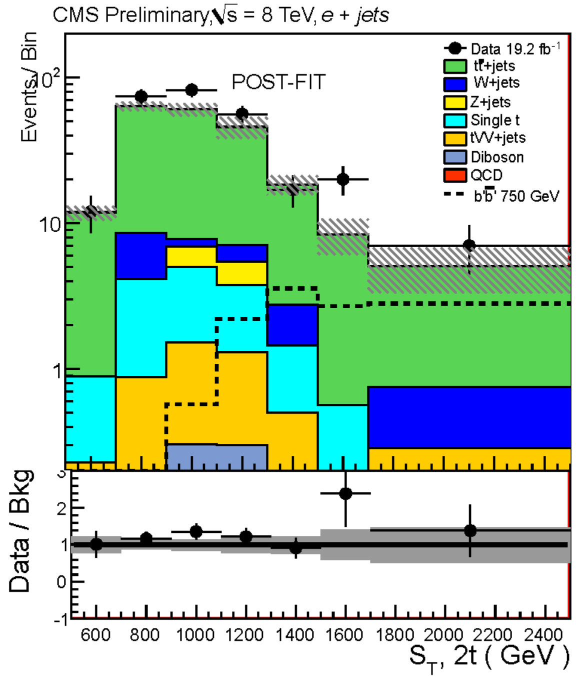
Jet Charge (k=0.3)

Search for b' pair production in l+jets

- Single lepton, at least 4 jets, one b-tag
- Again, focus on boosted (vector-boson-tagged) jets
- Only ST considered to be more model-independent
- Sensitive to tWtW, tWbZ, tWbH, bZbZ, bZbH, bHbH final states

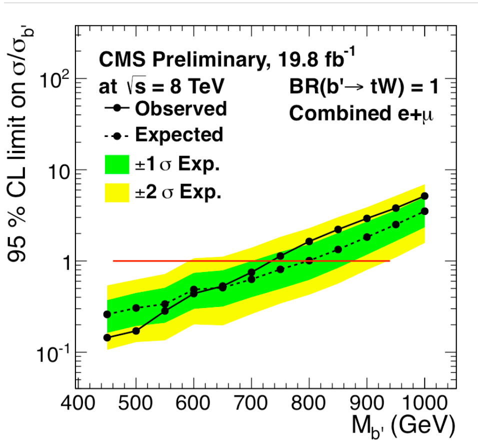


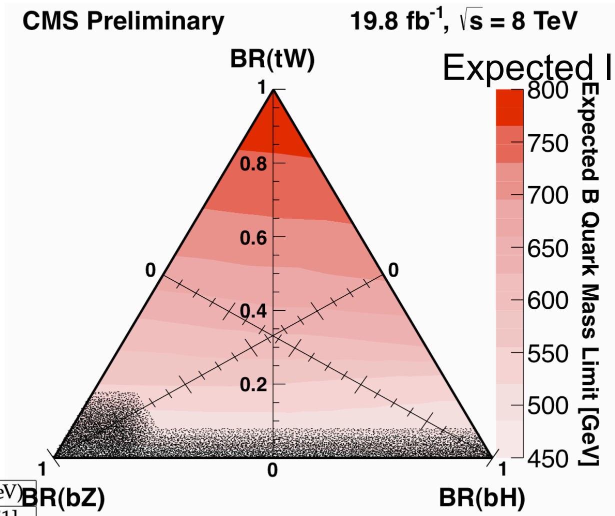




src: CMS PAS B2G-12-019

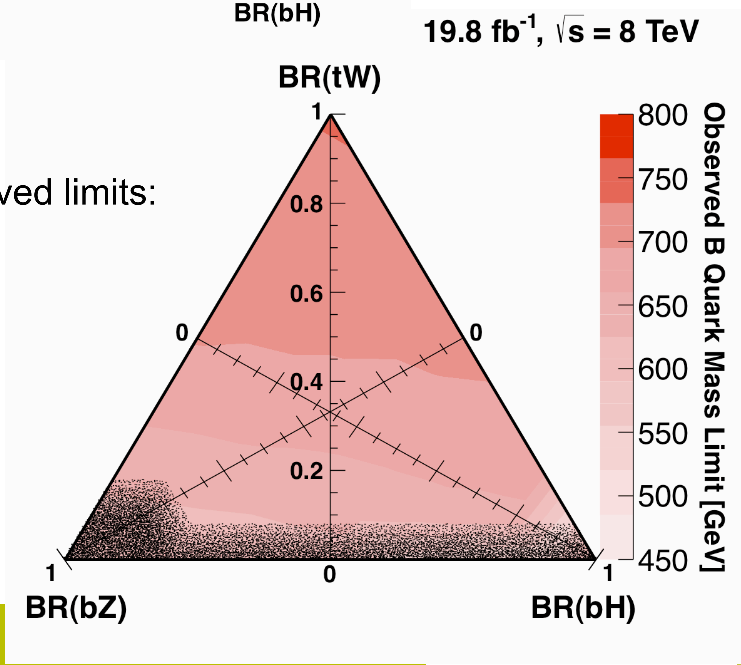
Combined result in BR plane: b'





Comb#	Br(tW)	Br(bH)	Br(bZ)	Obs (GeV)	Exp (GeV)	±1σ (GeV)	±2σ(GeV)
(0)	0.5	0.25	0.25	700	689	[586,782]	[516,851]
(1)	0.0	0.0	1.0	582	<450	[<450,495]	[<450,566]
(2)	0.0	0.2	0.8	601	<450	[<450,502]	[<450,565]
(3)	0.0	0.4	0.6	606	<450	[<450,506]	[<450,557]
(4)	0.0	0.6	0.4	546	<450	[<450,504]	[<450,530]
(5)	0.0	0.8	0.2	523	<450	[<450,501]	[<450,515]
(6)	0.0	1.0	0.0	634	454	[<450,514]	[<450,599]
(7)	0.2	0.0	0.8	639	555	[476,633]	[<450,721]
(8)	0.2	0.2	0.6	647	557	[484,645]	[<450,728]
(9)	0.2	0.4	0.4	653	555	[494,657]	[<450,732]
(10)	0.2	0.6	0.2	665	543	[500,664]	[<450,748]
(11)	0.2	0.8	0.0	671	532	[501,672]	[<450,755]
(12)	0.4	0.0	0.6	678	633	[557,741]	[473,810]
(13)	0.4	0.2	0.4	686	648	[560,745]	[486,814]
(14)	0.4	0.4	0.2	686	657	[562,754]	[493,825]
(15)	0.4	0.6	0.0	695	666	[555,760]	[503,830]
(16)	0.5	0.5	0.0	707	702	[594,792]	[522,861]
(17)	0.6	0.0	0.4	706	710	[599,802]	[537,866]
(18)	0.6	0.2	0.2	709	717	[609,810]	[542,875]
(19)	0.6	0.4	0.0	712	722	[625,816]	[544,879]
(20)	0.8	0.0	0.2	720	757	[666,845]	[567,905]
(21)	0.8	0.2	0.0	725	762	[671,852]	[574,908]
(22)	1.0	0.0	0.0	732	797	[700,873]	[589,934]

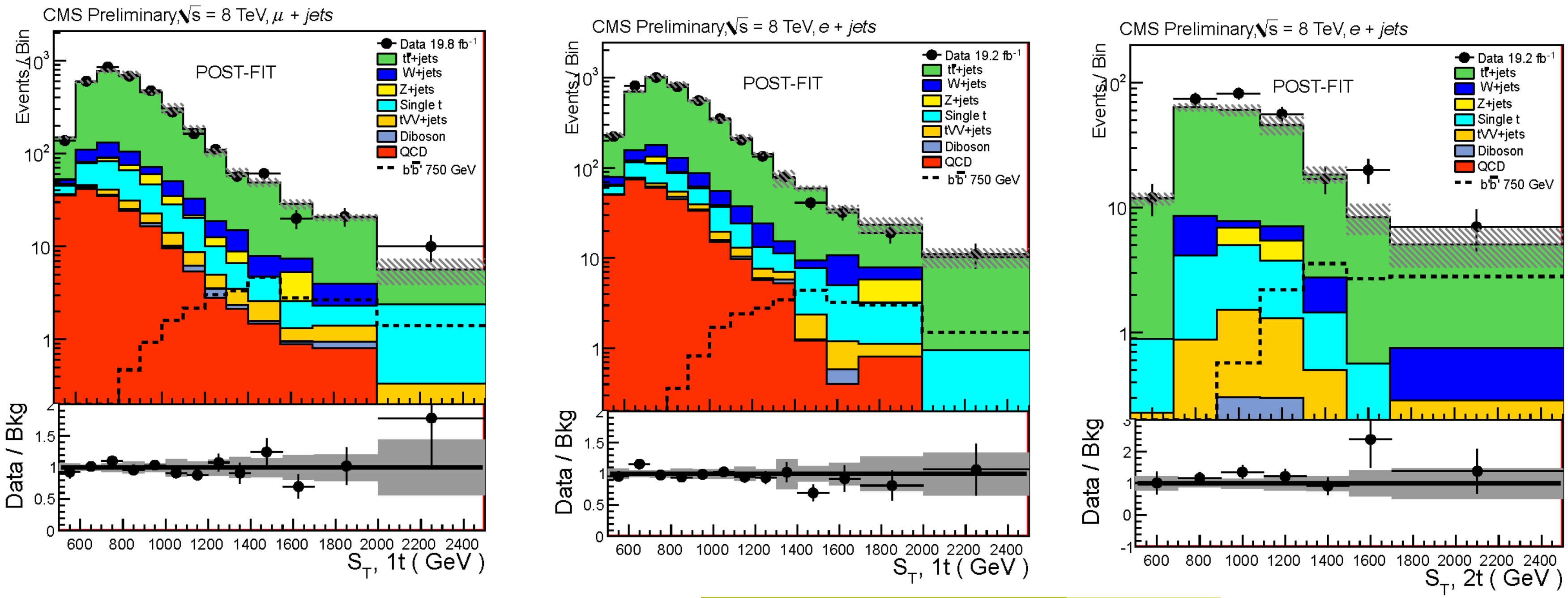
Observed limits:



src: CMS PAS B2G-12-019

Search for b' pair production in l+jets

- Single lepton, at least 4 jets, one b-tag
- Again, focus on boosted (vector-boson-tagged) jets
- Only ST considered to be more model-independent
- Sensitive to tWtW, tWbZ, tWbH, bZbZ, bZbH, bHbH final states



src: CMS PAS B2G-12-019