INFN and University of Milano Bicocca [1]

Search for a heavy resonances decaying in final states with boosted W/Z jet in CMS at 8 TeV

Speaker: **Raffaele Gerosa[1]**

on behalf of the CMS Collaboration

6th International Workshop on Boosted Object, UCL, London (UK)

Introduction

Many theories beyond SM predict existence of massive resonances coupled to vector bosons (**X → WW, WZ, ZZ**)

- **RS and Bulk Gravitons**
- **Extended Gauge models: W' → WZ**
- **Heavy scalars: SM higgs like, EWK singlet, 2HDM, psuedo-scalar Higgs**

In this talk, the following boosted results at 8 TeV are considered:

- **WW, WZ, ZZ semi-leptonic** → ***CMS-EXO-13-009, CERN-PH-EP-2014-076, arXiv:1405.3447 [hep-ex]***
- **WW, WZ, ZZ fully hadronic** → ***CMS-EXO-12-024, CERN-PH-EP-2014-071, arXiv:1405.1994 [hep-ex]***

NEW: combination X → VV analysis + model independent upper limit

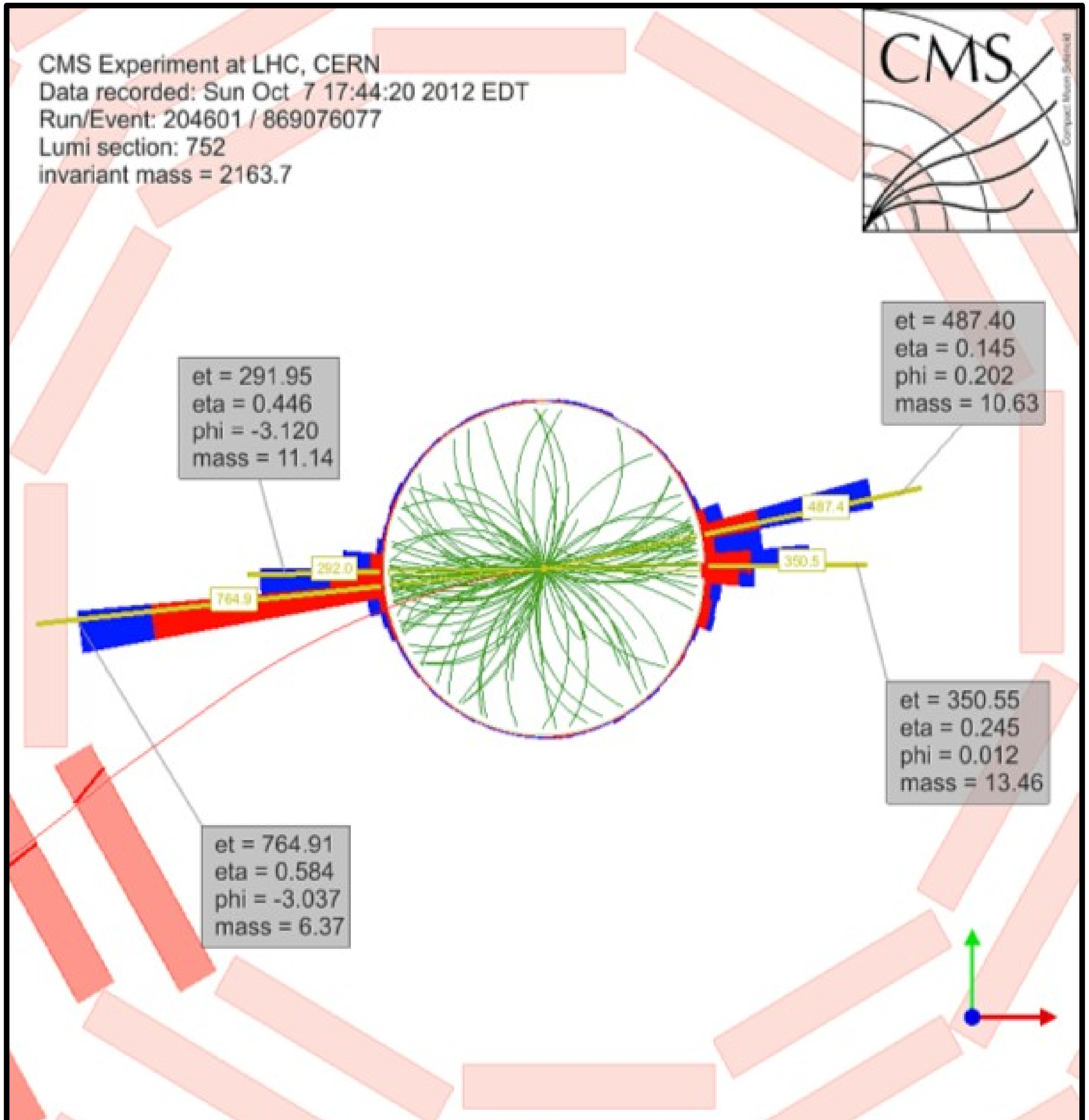
Additional CMS analysis using W/Z boson jet:

- **W/Z jet tagging in CMS** → ***CMS-PAS-JME-13-006, http://cds.cern.ch/record/1577417?ln=en***
- **Higgs like scalars decaying WW** → **CMS-PAS-HIG-13-008, https://cds.cern.ch/record/1546778**

See more on boosted analysis using **W/Z and Top jets in Kevin Nash** talk

https://indico.cern.ch/event/302395/session/11/contribution/20

X → VV → di-jet final state search



- Search performed at 8 TeV, ~20 fb^{-1}
- Heavy resonance search, **M > 1 TeV**
- **Hadronic W/Z bosons** are highly boosted → merged into a single big R-cone jet

✔ **Strategy: bump search in M_{jj}**

- **Signal benchmarks :**

1) q* → W/Z (excited quarks)

2) G_{RS} → WW/ZZ → V_T bosons

3) G_{bulk} → WW/ZZ → V_L bosons

4) W' → WZ (heavy partner of SM W)

$X \rightarrow VV \rightarrow$ di-jet search: event selection

Signals:

Simulated by JHUGEN, Pythia 6 and Herwig++

Narrow resonances for the chosen k/M_{pl} value

Backgrounds:

QCD multijet events:
Herwig++
MG+Pythia6

Signature:

$X \rightarrow WW/WZ/ZZ \rightarrow 2$ W/Z jets

Kinematics:

Two high p_T jets back-to-back

Online selection:

Two jet $p_T > 30$ GeV, $|\eta| < 2.5$, $|\Delta\eta_{jj}| < 1.3$, $M_{jj} > 890$ GeV

Offline selections:

Two **leading** CA8 jet considered as **W/Z candidates**

1) pruned mass [70,100] GeV

2) N-subjettiness ratio τ_2/τ_1 used to classify events:

$\tau_2/\tau_1 < 0.5 \rightarrow$ **high purity** τ_2/τ_1 in [0.5,0.75] $\rightarrow$ **low purity**

Merged Jet

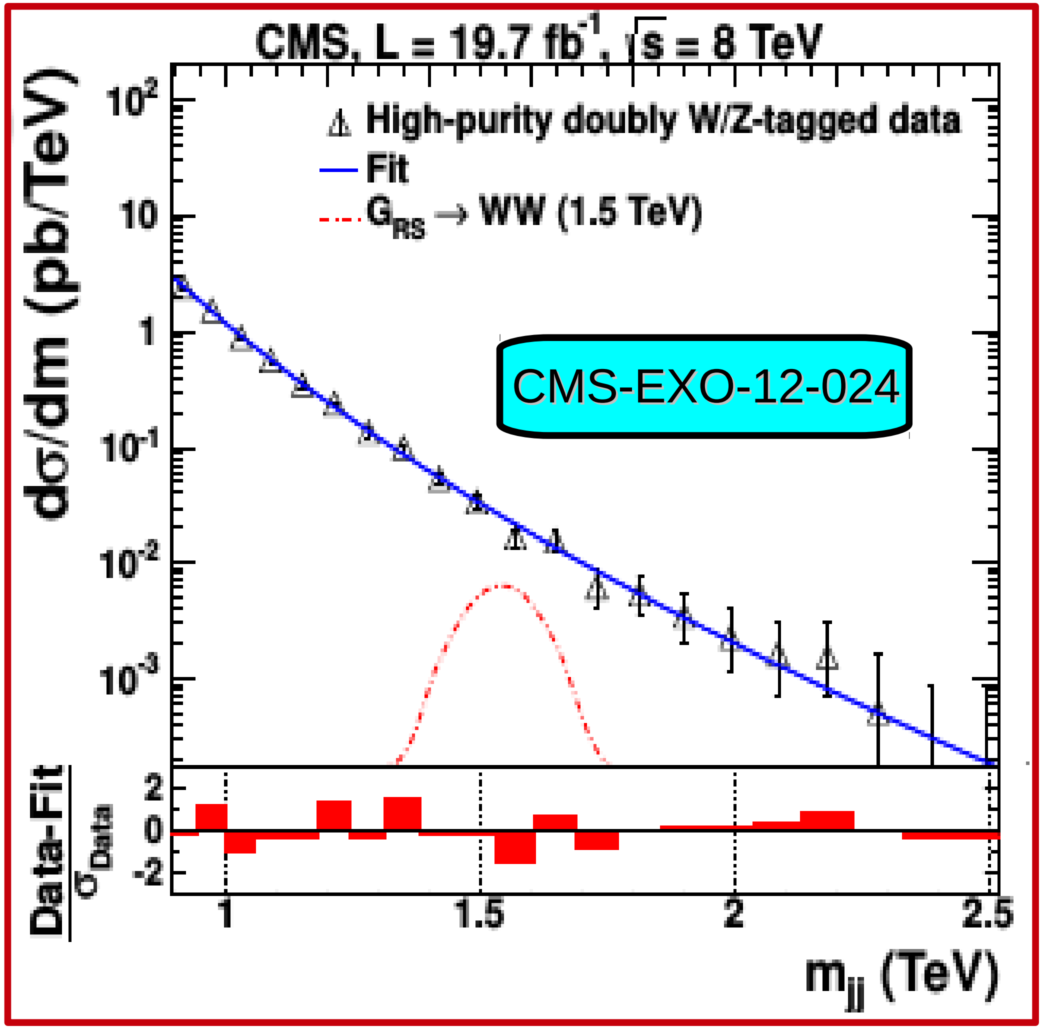
Use jet substructure used to identify single jets containing the decay products of the hadronic W/Z

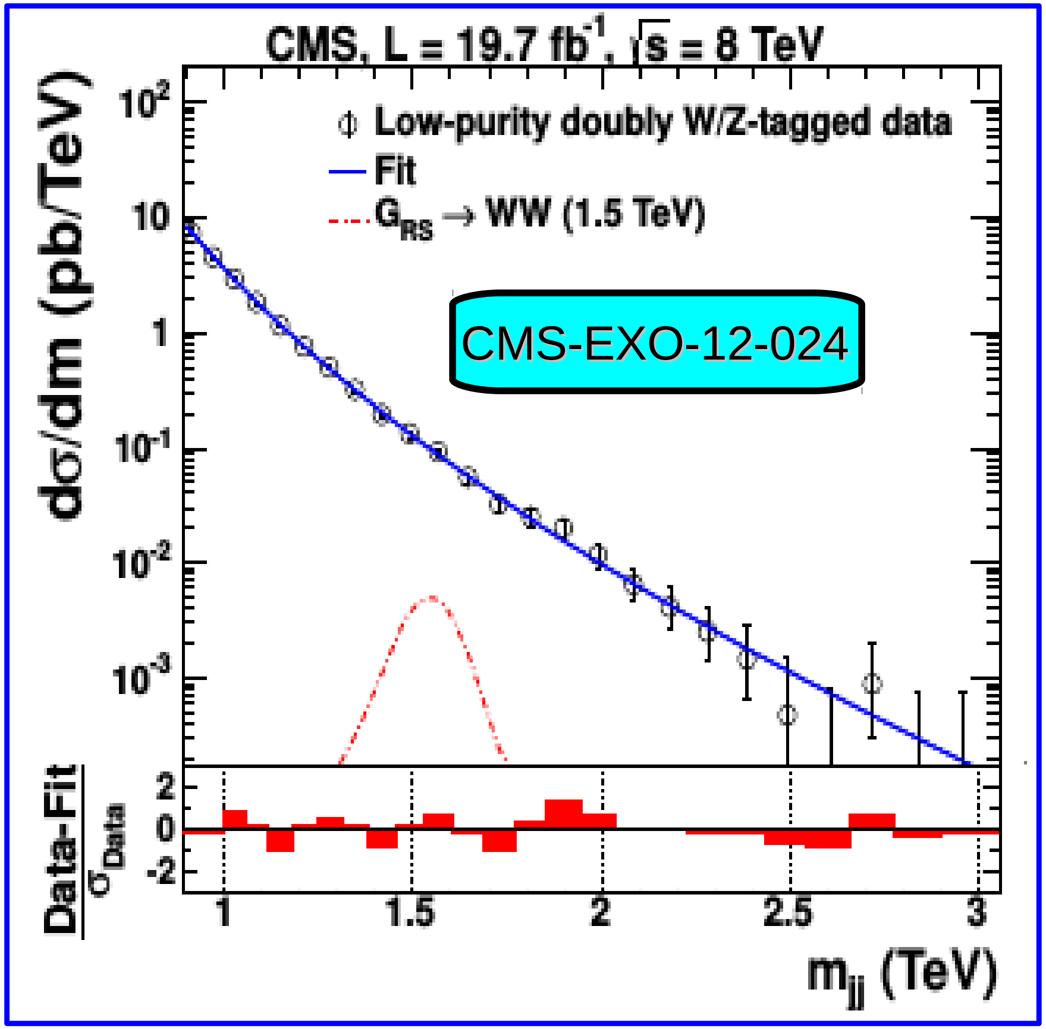
Data-Driven

QCD multijet background extracted from data. Simulation used only as a cross check

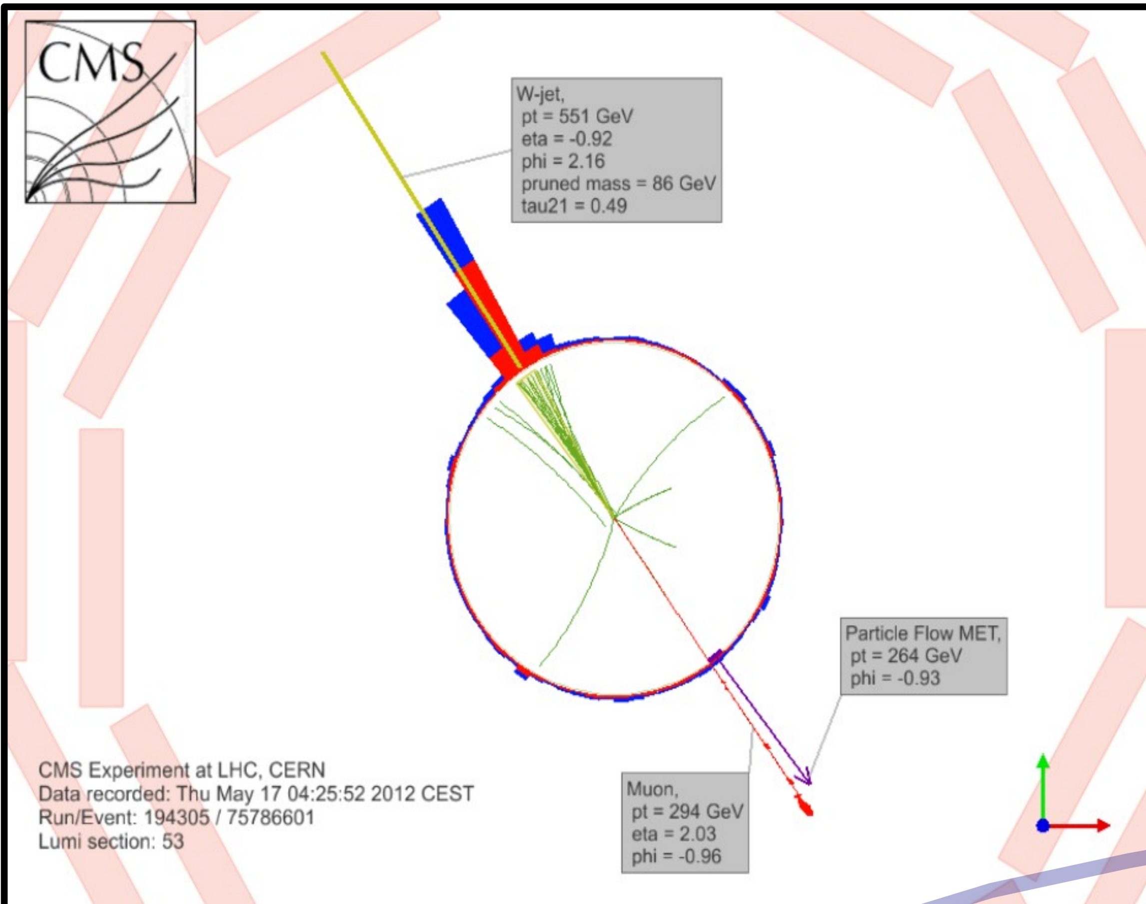
X→VV→di-jet search: mass spectrum

- **Background** → extracted from data by means of a smoothing test



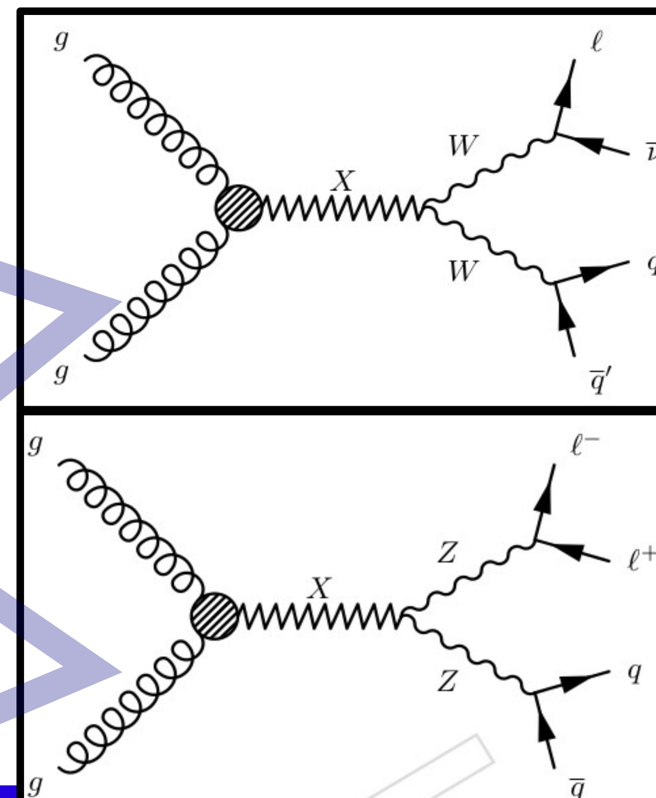


X → VV → semi-leptonic final state



- Heavy resonance search, **M > 0.6 TeV**
- Semi-leptonic WW, ZZ final states (μ,e)
- **Hadronic W/Z bosons** are highly boosted → merged into a single big R-cone jet

✔ **Strategy:** bump search in M_{VV}



- **Signal benchmark :**

G_{bulk} → WW/ZZ → V_L bosons

theoretical cross section revisited → 4 time smaller → effect on exclusion

$X \rightarrow VV \rightarrow$ semi-leptonic search: strategy

Signals:

Simulated by JHUGEN and Pythia 6

Backgrounds:

W+jet: MG+Pythia6
Herwig++

DY+jet: MG+Pythia6

$t\bar{t}$: Powheg+Pythia6
mc@nlo+Herwig++

VV: Pythia6

Signature:

$X \rightarrow WW \rightarrow l(\mu,e)$ + MET + W-jet

$X \rightarrow ZZ \rightarrow 2l(\mu,e)$ + Z-jet

Merged Jet

Use jet substructure to identify single jets containing decay products of hadronic W/Z

Kinematics:

Two high p_T W or Z bosons back-to-back

Online selection:

Single lepton high p_T triggers, no isolation: $p_T > 40$ (80) GeV μ (e)

Di-lepton triggers, no isolation: $p_T > 22$ (8) GeV $\mu\mu$, $p_T > 33$ GeV ee

Data-Driven

W+jet / DY+jet background extracted from data in the jet mass sideband

Offline selections:

Two or one isolated lepton (μ,e)

Leading CA8 jet taken as W/Z-jet candidate

W/Z jet tagging is consistent with di-jet analysis
(full description in the backup)

X→VV→semi-leptonic : background

After analysis selections, **SM V+jets are main backgrounds**

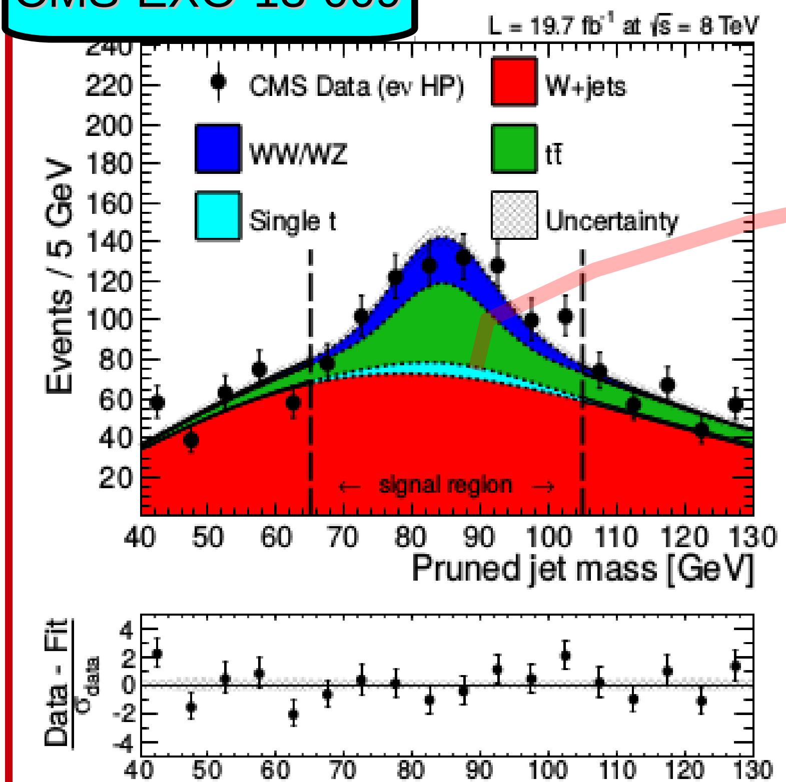
Minor contribution are taken from simulation + corrections from control regions

V-jets background is estimated from data using jet mass sideband control region

W+jets SB: m_J [40,65] U [105,130] GeV **DY+jets SB**: m_J [40,70] U [110,130] GeV

Overall normalization in signal region from jet mass data fit in the sidebands

CMS-EXO-13-009



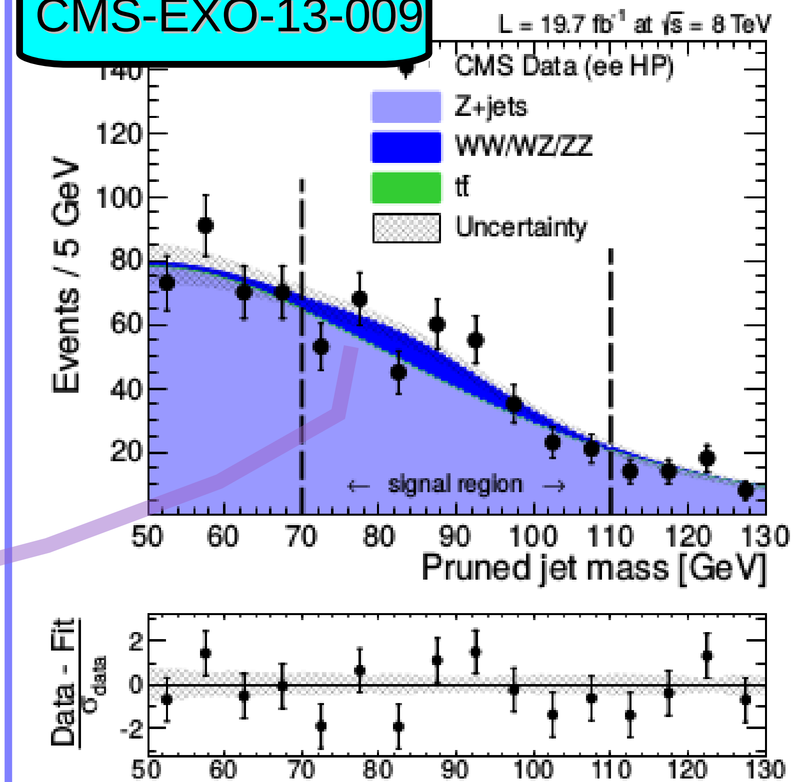
Resonant contamination from VV and tt

Low SB dominated by W+jets

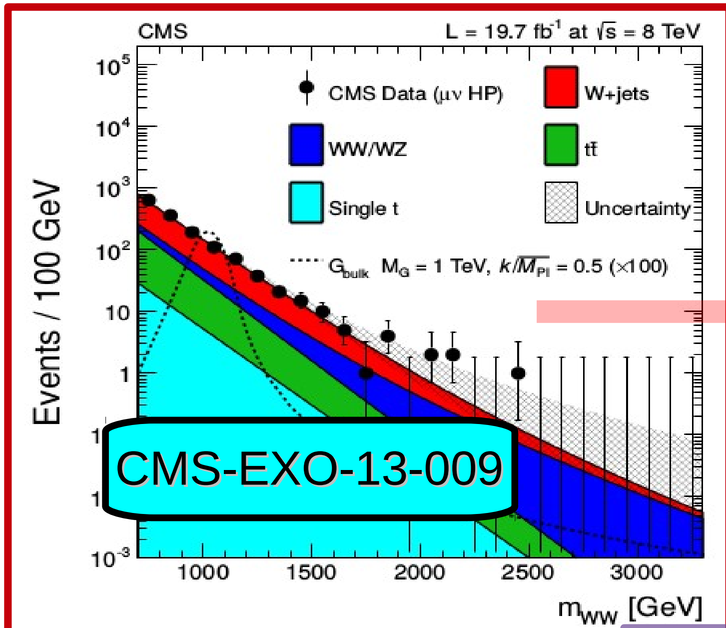
Resonant contamination from VV

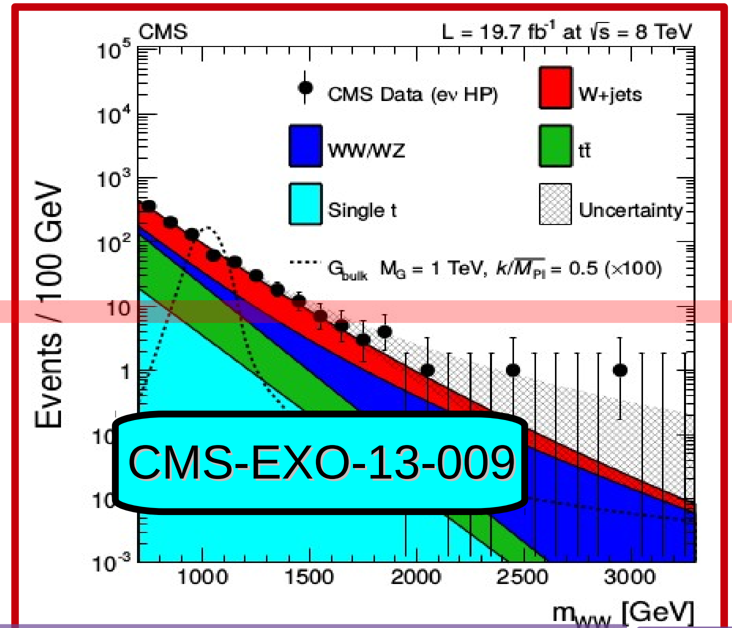
Low SB dominated by Z+jets

CMS-EXO-13-009



$X \rightarrow VV \rightarrow$ semi-leptonic : M_{VV} spectrum





M_{VV} spectrum for l+ν+jet

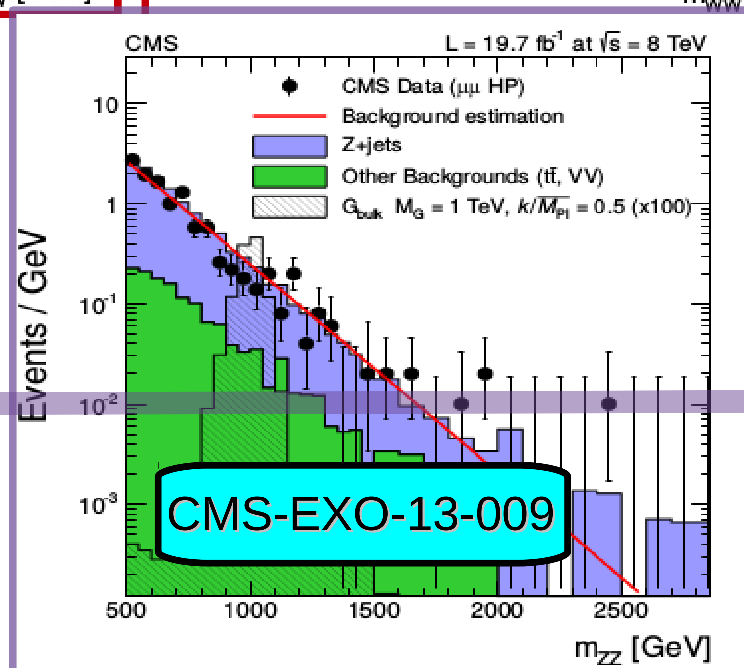
Muon HP (left)

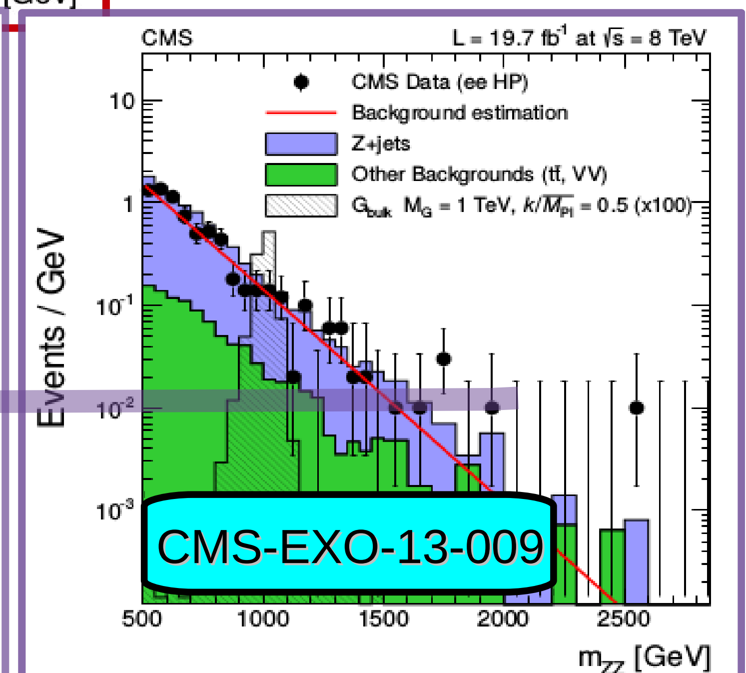
Electron HP (right)

M_{VV} spectrum for l+l+jet

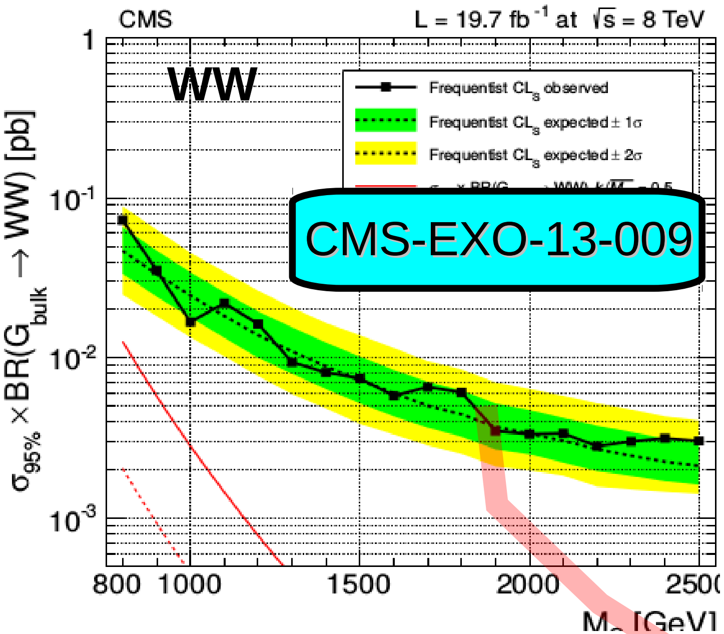
Muon HP (left)

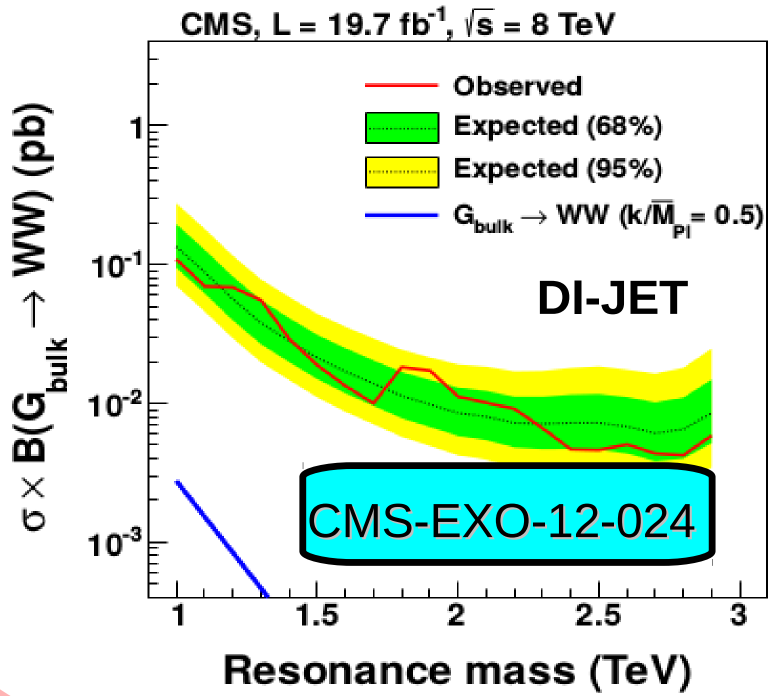
Electron HP (right)

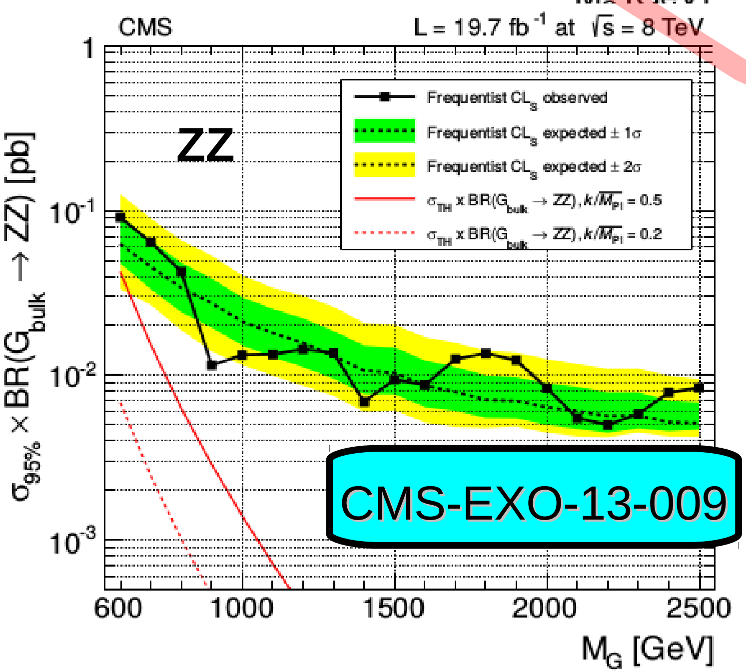




X→VV searches : narrow bulk graviton







Results:

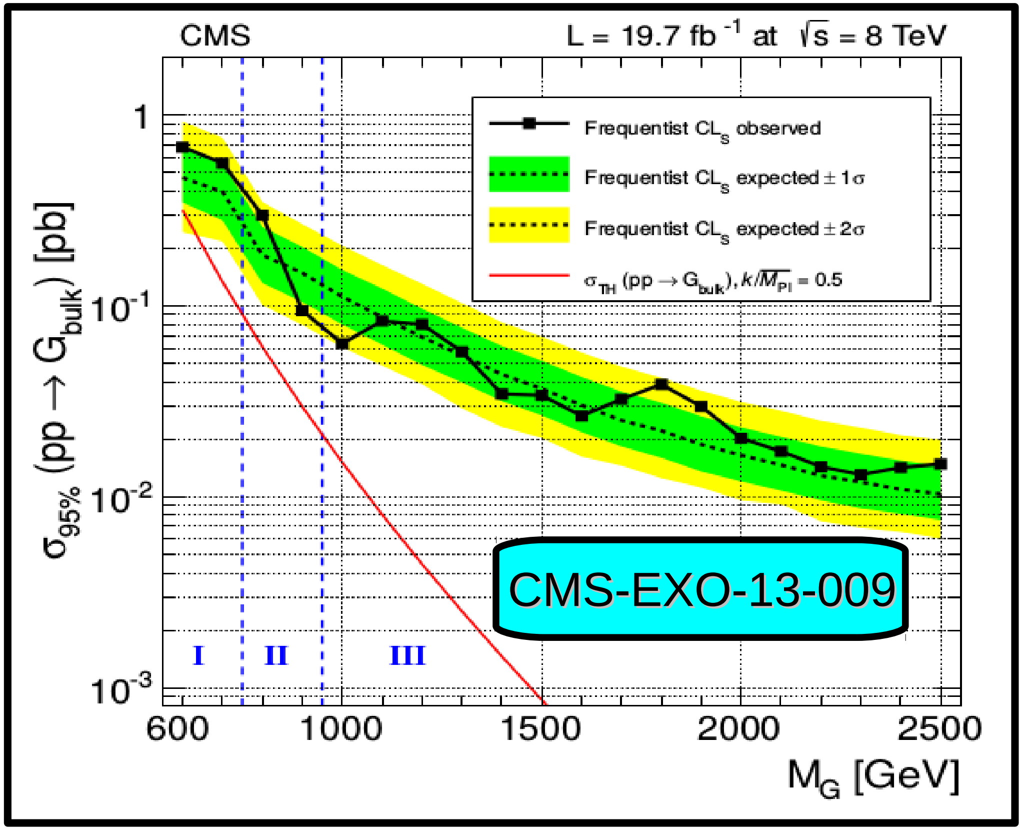
WW → l+ν+jet is the most sensitive on the whole explored mass range

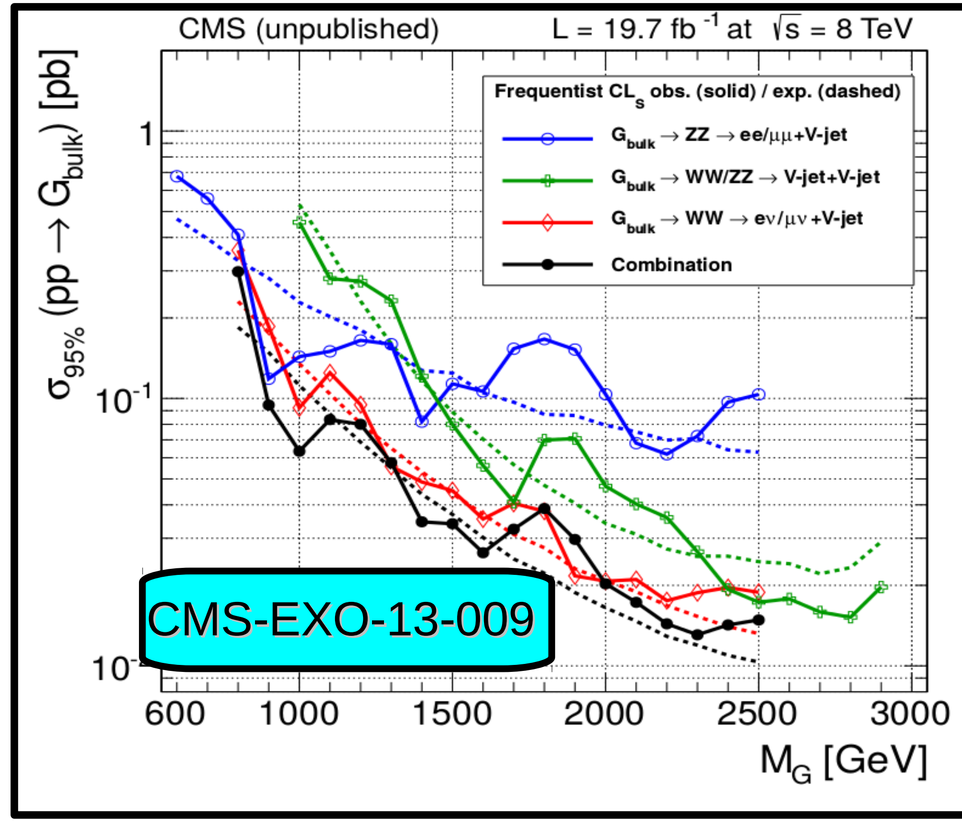
G_{bulk} → VV not excluded in any of the exclusive final state search for $k/M_{pl} < 0.5$

Results for q*, W' and G_{RS} in the backup

G_{bulk} → VV combination

- **Combination:** l+ν+jet , l+l+jet and fully hadronic VV searches are combined together
- **Correlated sys:** V-tagging, luminosity, jet scale and resolution, lepton scale and resolution





Results:

l+ν+jet dominates in the range [800,2500] GeV → gain ~20% with the combination

No exclusion for k / M_{pl} < 0.5

X → VV : model independent limit

Reinterpretation: simplify the analysis → **drop LP** category and **merge lepton flavour**

Provide upper limits in terms of observed number of signal events

Reconstruction + selection efficiency for vector boson is evaluated

In this way, we can estimate excluded number of signal events for a generic model

Generic model is not restricted to narrow signals → make a scan vs M_X and Γ_X

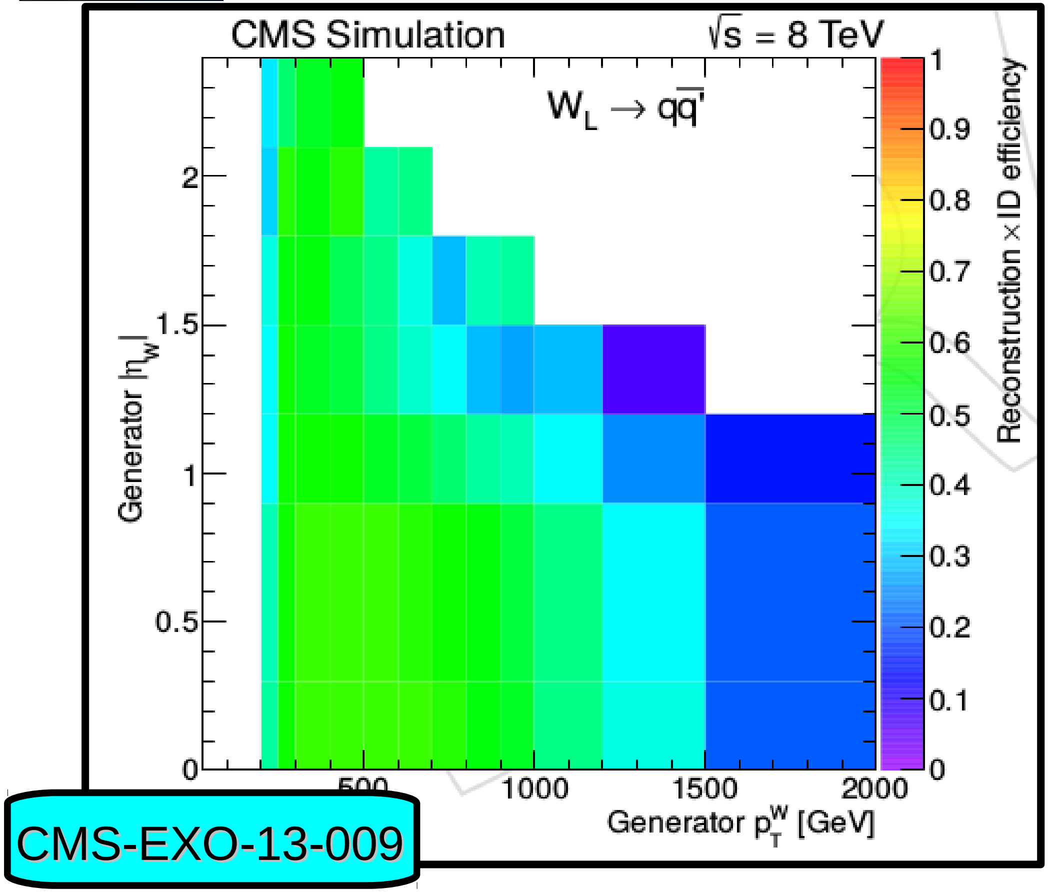
→ **intrinsic line-shape** supposed to be a **Breit-Wigner** with Γ_X

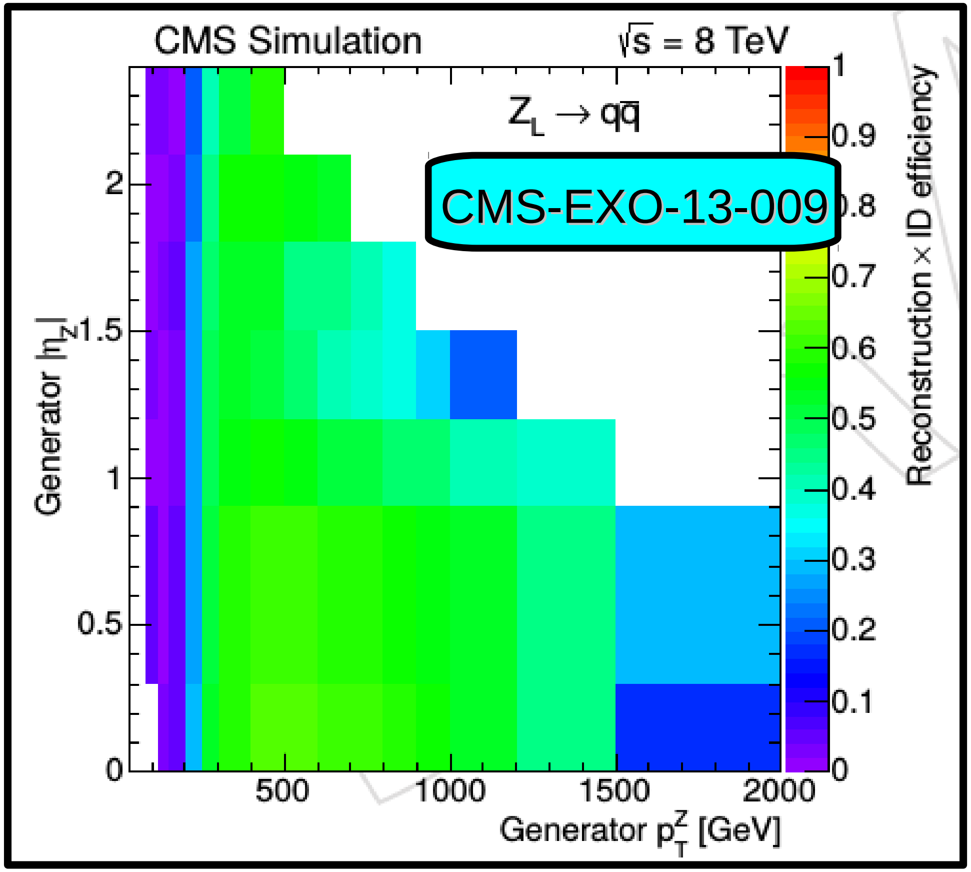
→ **Resulting signal shape is a convolution between BW and resolution**

- **Vector boson efficiency** →provided as a function of V-kinematics (p_T,η) and V-polarization

1) Pre-select signal events applying all the acceptance cuts of the analysis (gen level)
2) Reconstruction + identification efficiencies are evaluated for each vector boson in the event in a independent way
3) All re-weighting applied to signal events are included in the efficiency calculation

$X \to VV$: efficiency maps





- Be careful in cases where V is transverse polarized → **RS1 Graviton** from **MG+Pythia6** used

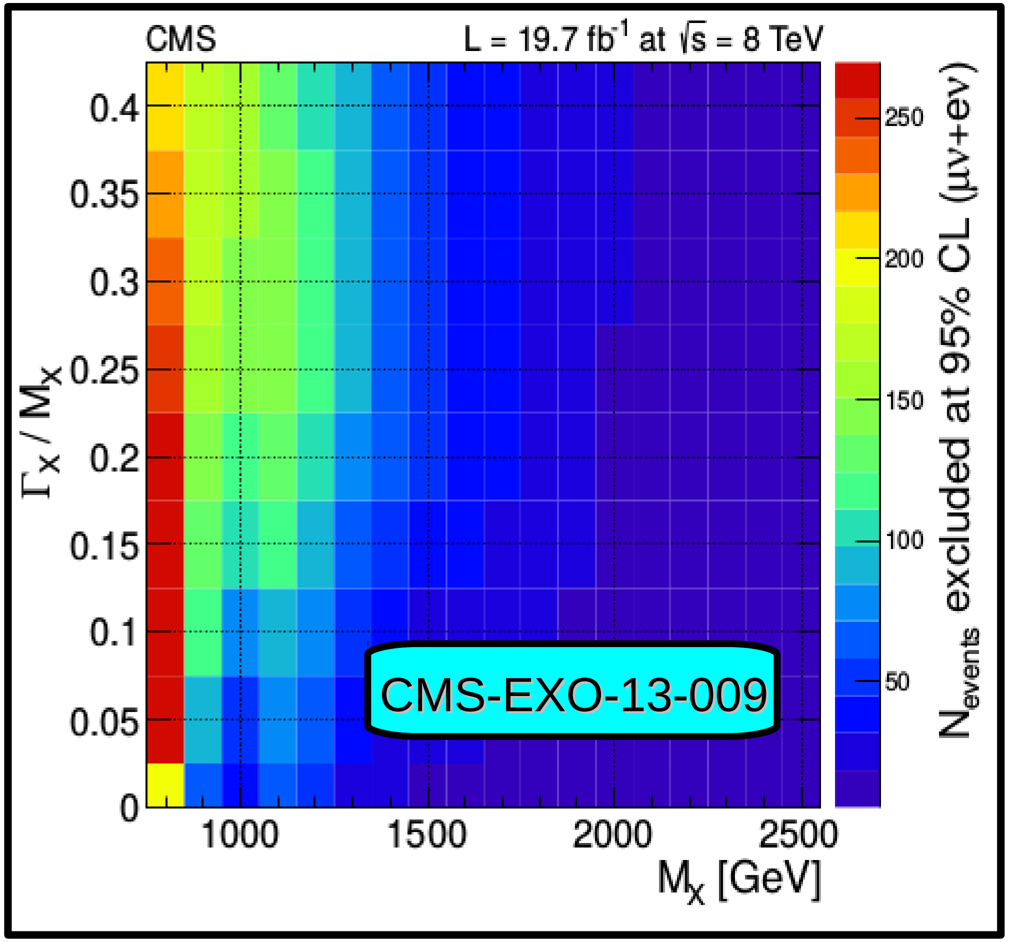
Loss in the efficiency due to the vector boson polarization found to be independent of p_T and η → 0.85 k-factor from V_L to V_T

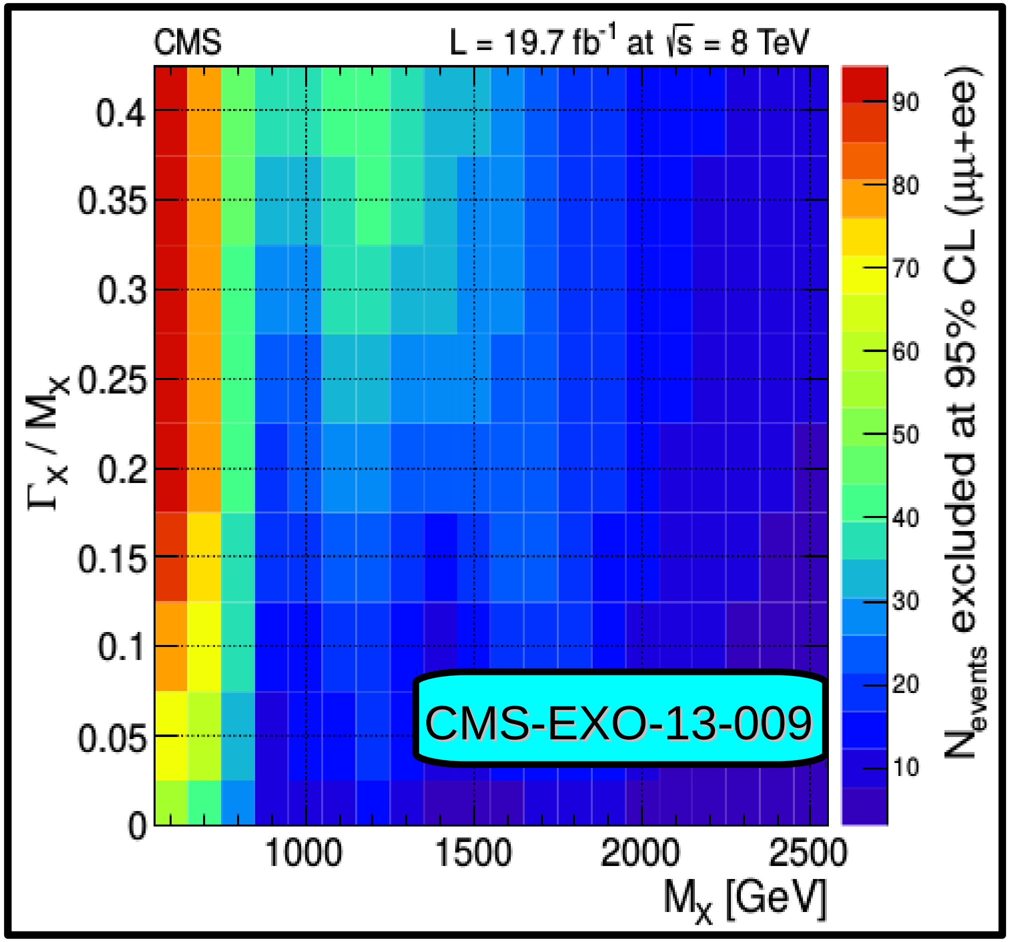
- **Correlations among physics objects not considered → 15% extra systematics**

All the efficiency maps are in the backup

X → VV : model independent result

- l+ν+jet and l+l+jet not combined in order to avoid assumption on BR(G→WW), BR(G→ZZ)
- **Degradation of performance vs Γ_x** ; large resonances fluctuation due to very broad shape





X → VV searches : conclusions

- Searches for X → VV in fully hadronic and semi-leptonic have been performed at 8 TeV
- Tools for tagging boosted W/Z bosons decaying hadronically have been widely used
- **Combination among all the X → VV searches in CMS for bulk graviton → no exclusion**
- **Model independent re-interpretations of these results is now possible**

Simplified upper limits provided in terms of excluded number of signal events **vs M_X and Γ_X**

Reconstruction + Selection efficiencies per **Vector Boson** provided in **(p_T,η)** tables

For **transverse polarized vector boson** apply a flat k-factor of **0.85**

- **No excesses found at 8 TeV ….. stay tuned for the next Run with 13 TeV data !**

Performance in Run II: introduction

- Use CMS full-simulation @13 TeV, in time average pileup <PU> = 40

- **Observables:** AK8 jets + CHS with p_T in [475,600] GeV used in this study

  Trimming: k_T sub-jets, $R_{filt} = 0.1$ and $p_T^{frac} > 3\%$

  Pruning: CA with R = 0.8 used for pruning, $z_{cut} = 0.1$ and $D_{cut} = 0.5$

  N-subjettiness (τ_1,τ_2,τ_2/τ_1) and Q-jet volatility

→ **Soft-Drop:** soft threshold $z_{cut} = 0.1$ and β = [-1,0,1,2]

→ **CMS Q/G Likelihood:** applied on both pruned jet and two leading pruned sub-jets

→ **Constituent subtracted mass:** pileup subtraction at single particle level

→ **$C_2(\beta)$:** double ratio of energy correlation function with β = [0,0.5,1,2]

Groomed masses are corrected for pileup by means of **safe 4V subtraction**

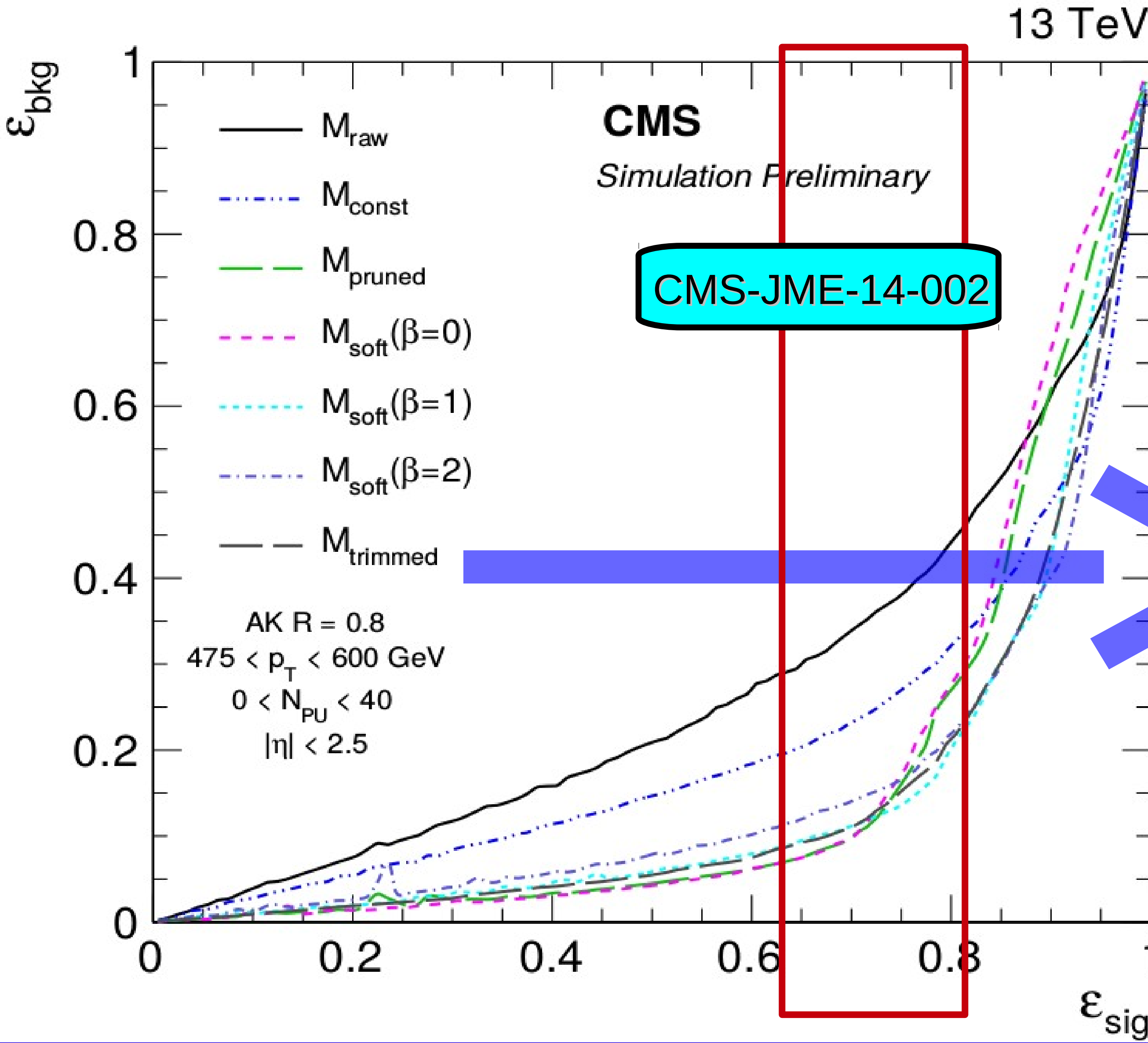
More on observables, data/MC, performance and correlation @8TeV in **Tobias** talk
https://indico.cern.ch/event/302395/session/17/contribution/32

More on high p_T performance for Run II in **Andreas** talk
https://indico.cern.ch/event/302395/session/10/contribution/18

V Tagging observables and correlations → ***CMS-JME-14-002,***
https://twiki.cern.ch/twiki/bin/viewauth/CMS/PhysicsResultsJME14002Draft

Mass variables performance

- Train a **single variable BDT**, only **leading jet** with p_T [475,600] GeV are considered
- **Signal** are W jets from RS Graviton decay, **background** from **QCD Pythia8**



Groomed mass compared w.r.t. raw one

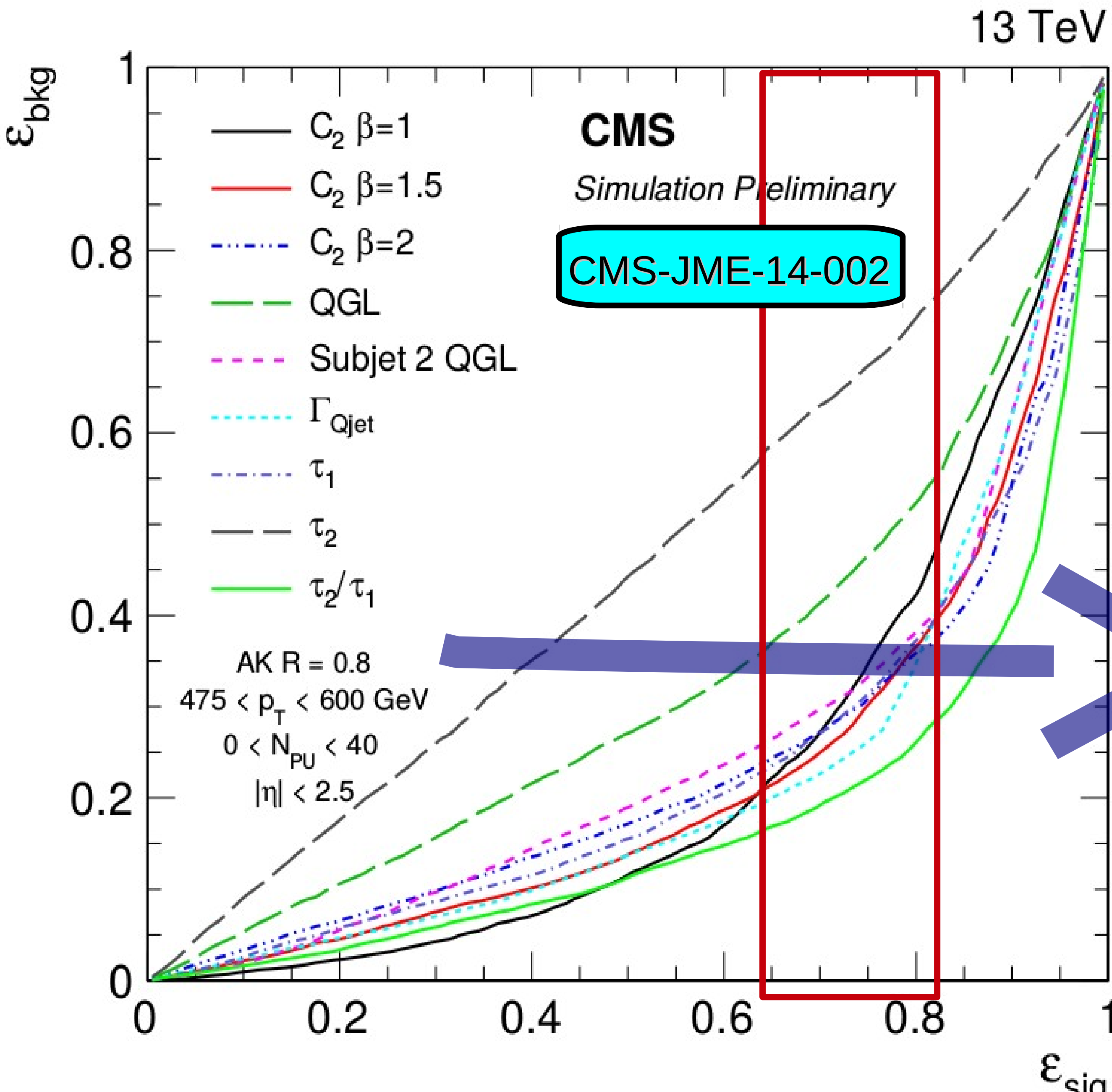
Best performing

Interesting to look at ε_{sig} [70,80] %

Trimming,Pruning
Soft drop $\beta = 1$

Pileup dependence ?

substructure variables performance



Many sub-structure observables

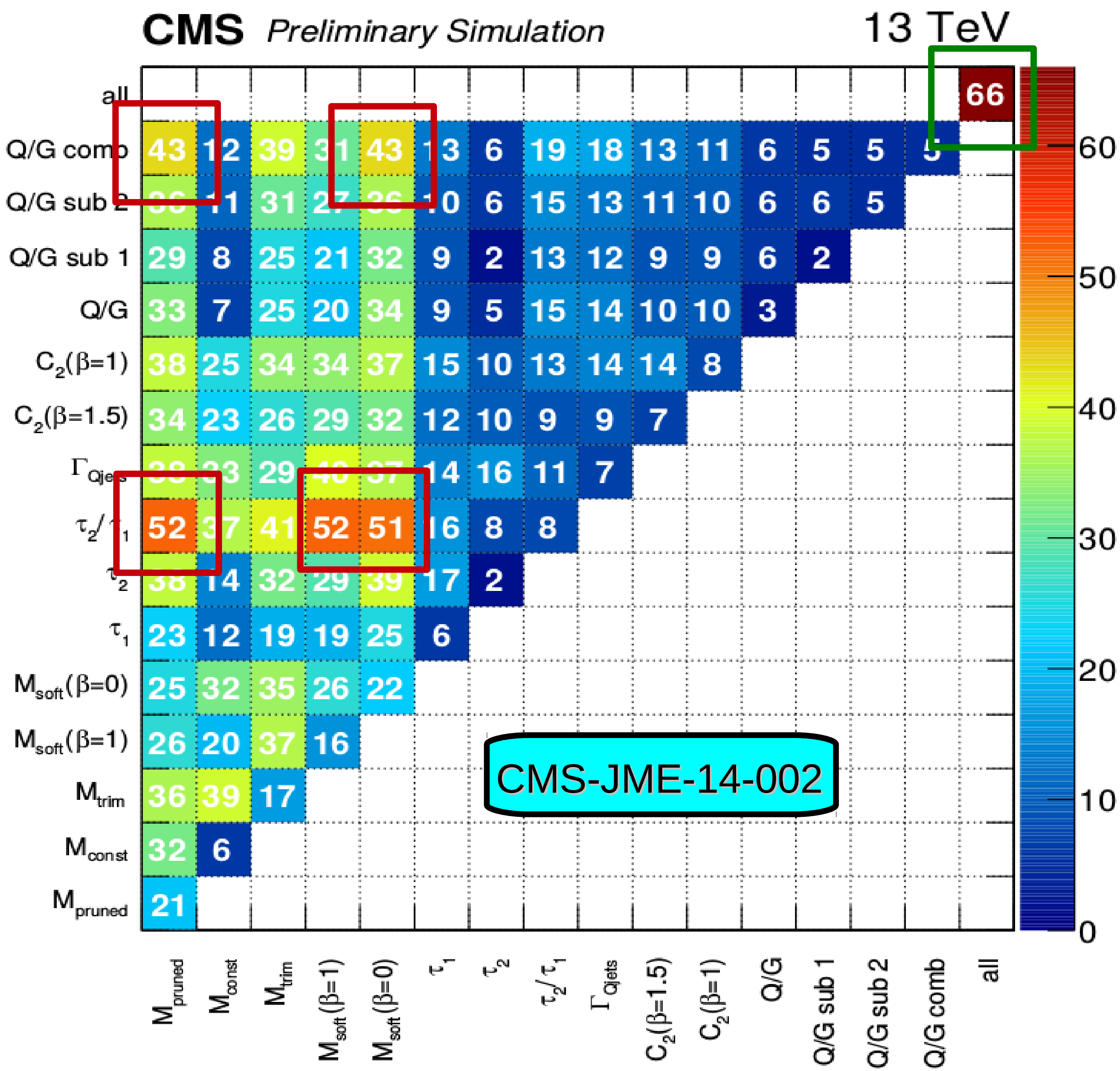
Best performing

Interesting to look at ε_{sig} [70,80] %

τ_2/τ_1, Qjet, $C_2(\beta = 1.5)$, Q/G Like on sub-jet 2

Pileup dependence ?

BDT performance: pairs



- Try to improve the single variable performance by means of **pairs combination**
- Do this in two independent pileup bins

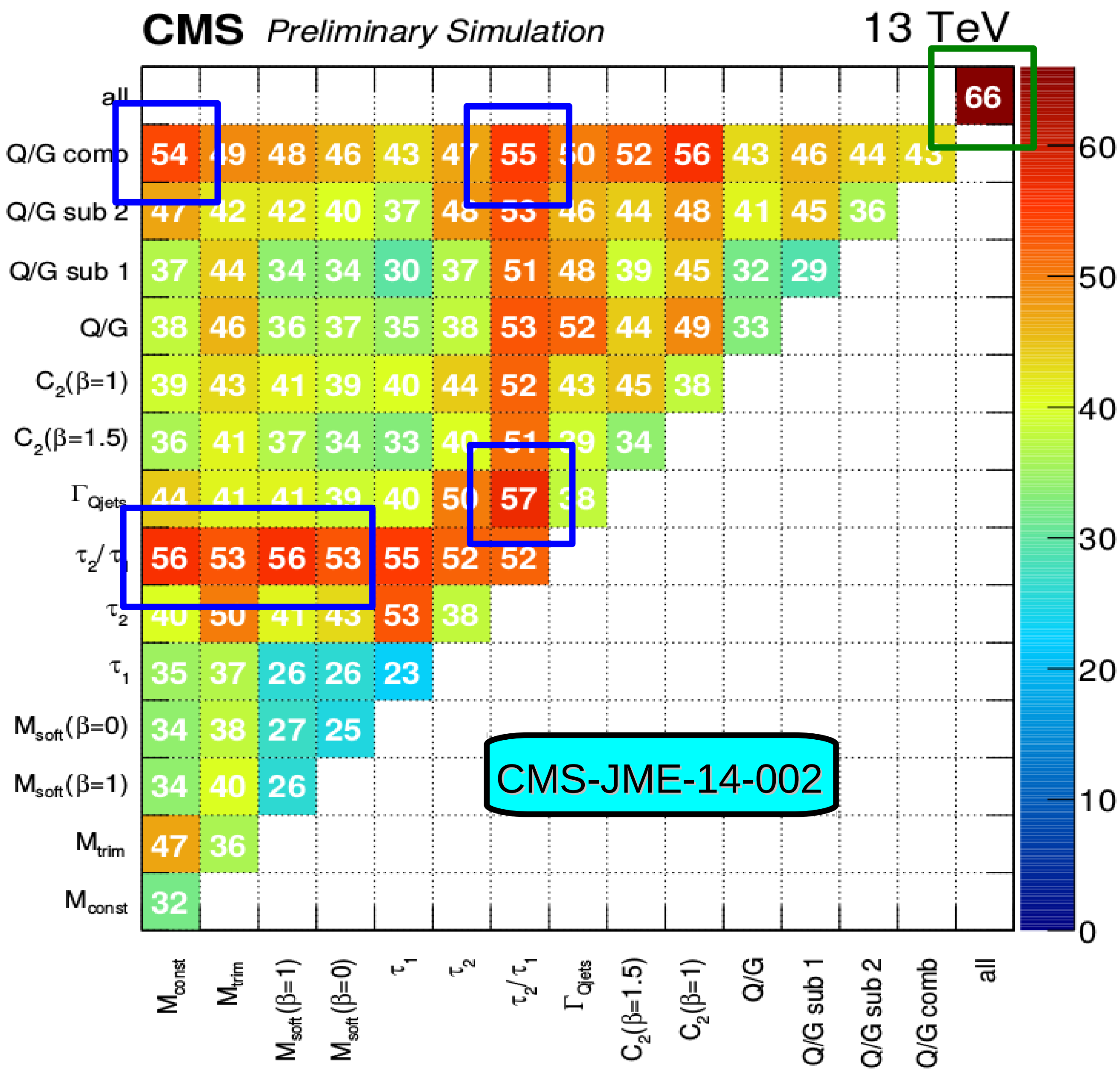
Best Performance

Performance is estimated as
$1/\varepsilon_{back}$ **for** $\varepsilon_{sig} = 50\ \%$

- **Pair → pruned** or **soft drop** mass combined with τ_2/τ_1 or **subjet Q/G combination**

Performance still far from what you get combining all the observables

BDT performance: triplets



- Try to improve the single variable performance by means of adding **pairs of variables to pruned mass**
- Do this in two independent pileup bins

Best Performance

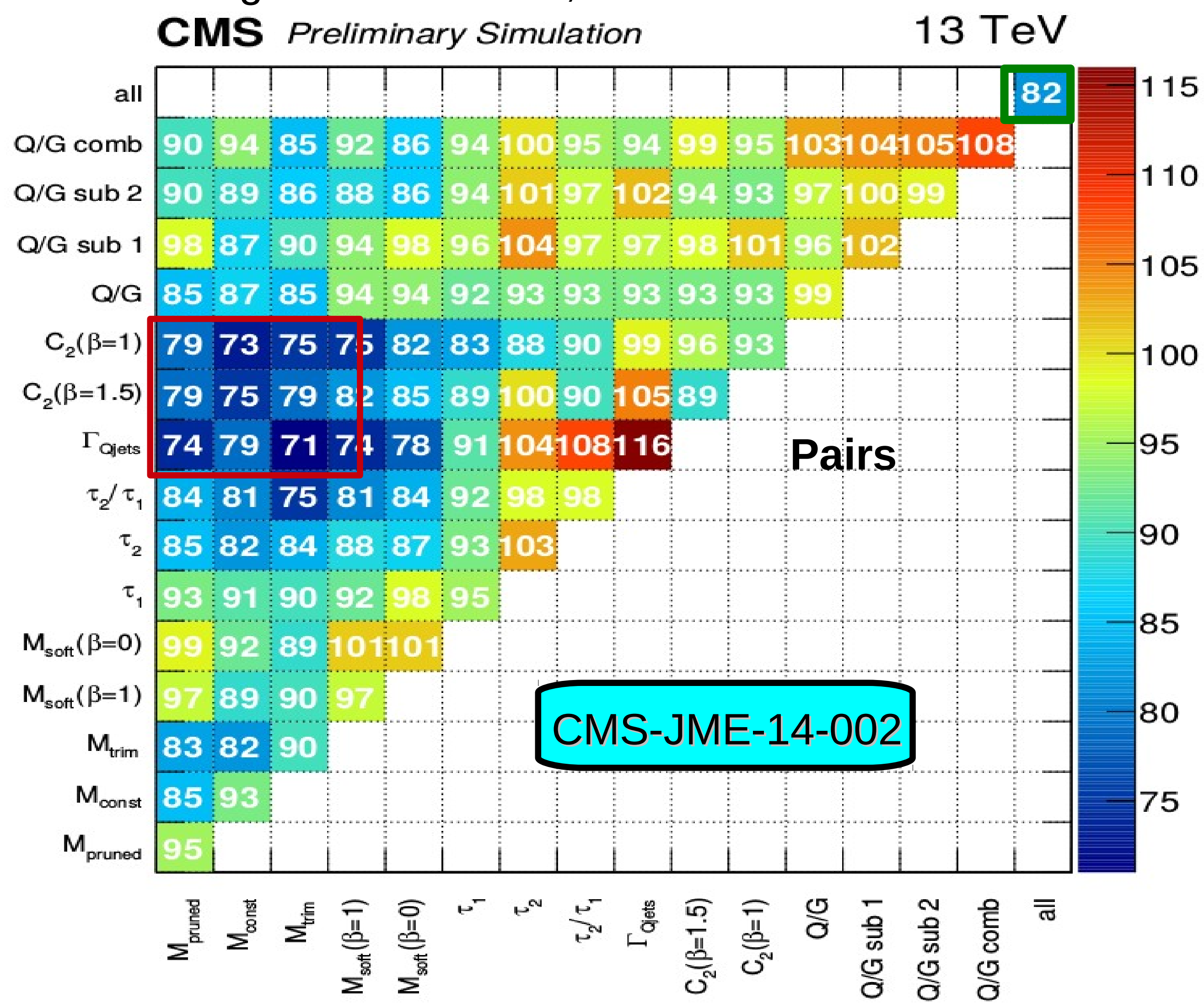
Performance is estimated as

$\mathbf{1/\varepsilon_{back}}$ **for** $\mathbf{\varepsilon_{sig} = 50\ \%}$

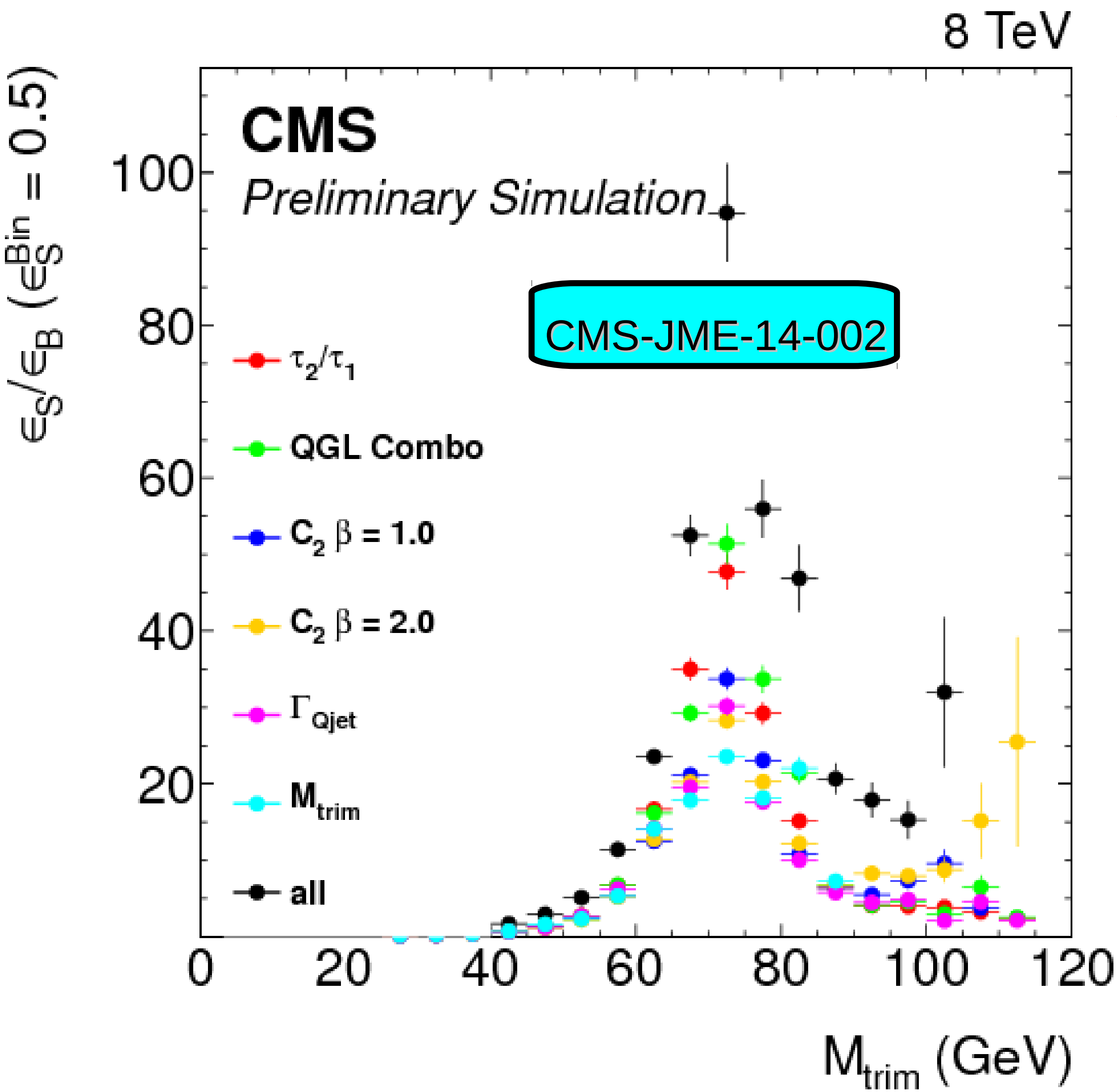
- **Triplets** → adding a **groomed mass** or **Qjets** or **Q/G likelihood to the best pair** can help in the performance
- Performance getting closer to the total combination in a BDT

BDT performance vs pile-up

- **BDT performance vs Pileup → percent ratio** of the previous maps
- Small effect is visible at single variable level; **20%** effect on the total combination



BDT Performance: summary plot



- Binned the events as a function of trimmed mass
- Consider a set of pairwise BDT combining trimmed mass + another variable
- In each M_{trim} bin, make a ROC curve for the BDTs, taking the background efficiency for ϵ_{sig} = 50%
- **QGL** and τ_2/τ_1 are adding performance in the peaking part

Performance in Run II: conclusions

- **First look at V-tagging performance @13TeV with full CMS simulation**

- **Soft drop mass** performs like pruning and trimming → many configuration of parameters have been tested

- New substructure observables have been tested:
  → **Q/G likelihood** on pruned jet and sub-jet → **looks promising !!**

  → **double ratio of ECF** can help in boost the V-tagging performance

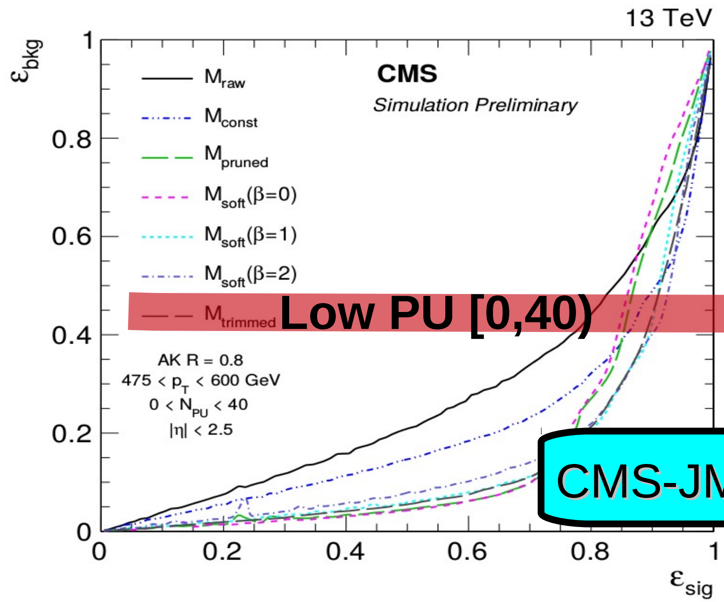
- **Pile-up dependence** looks under control for the investigated p_T regime:
  → worsening in the performance between **10-20 %** when observables are combined

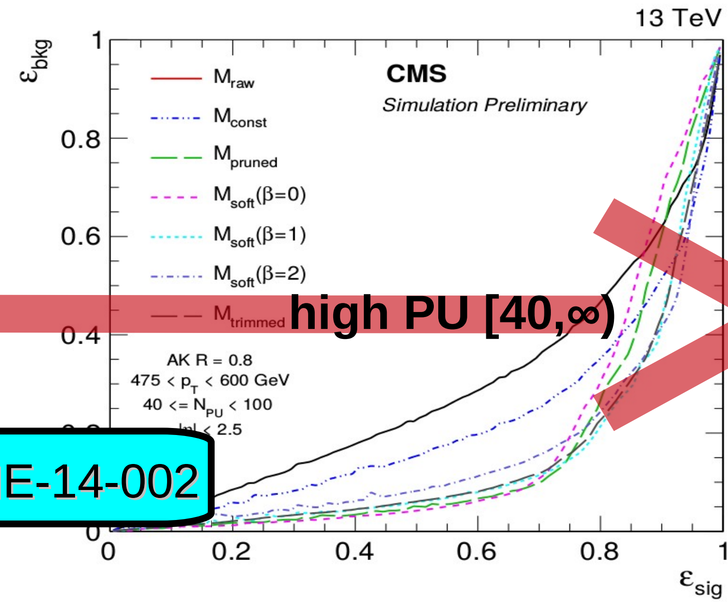
  → The bigger effect is visible on **Qjets volatility** and **C2(β)**

Backup slides

Single variable performances vs PU

- Train a **single variable BDT**, only **leading jet** with p_T [475,600] GeV is used
- **Signal benchmark** are W jets from RS Graviton (JHUGen), **background** from **QCD Pythia8**





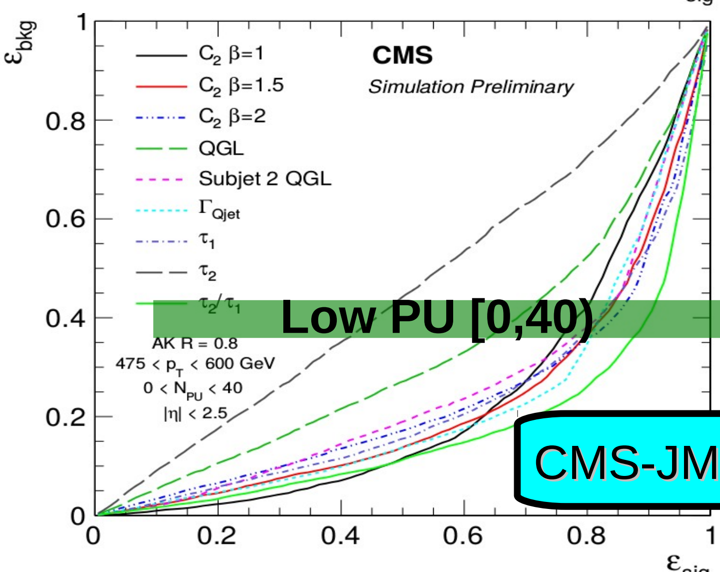
Groomed mass compared w.r.t. raw one

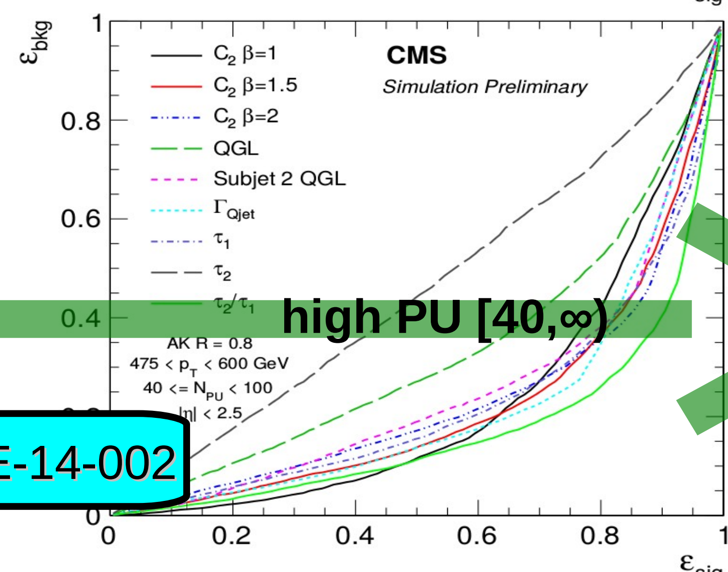
Best performing

Interesting to look at ε_{sig} [70,80] %

Trimming,Pruning
Soft drop $\beta = 1$

Pileup dependence??





Many sub-structure observables

Best performing

Interesting to look at ε_{sig} [70,80] %

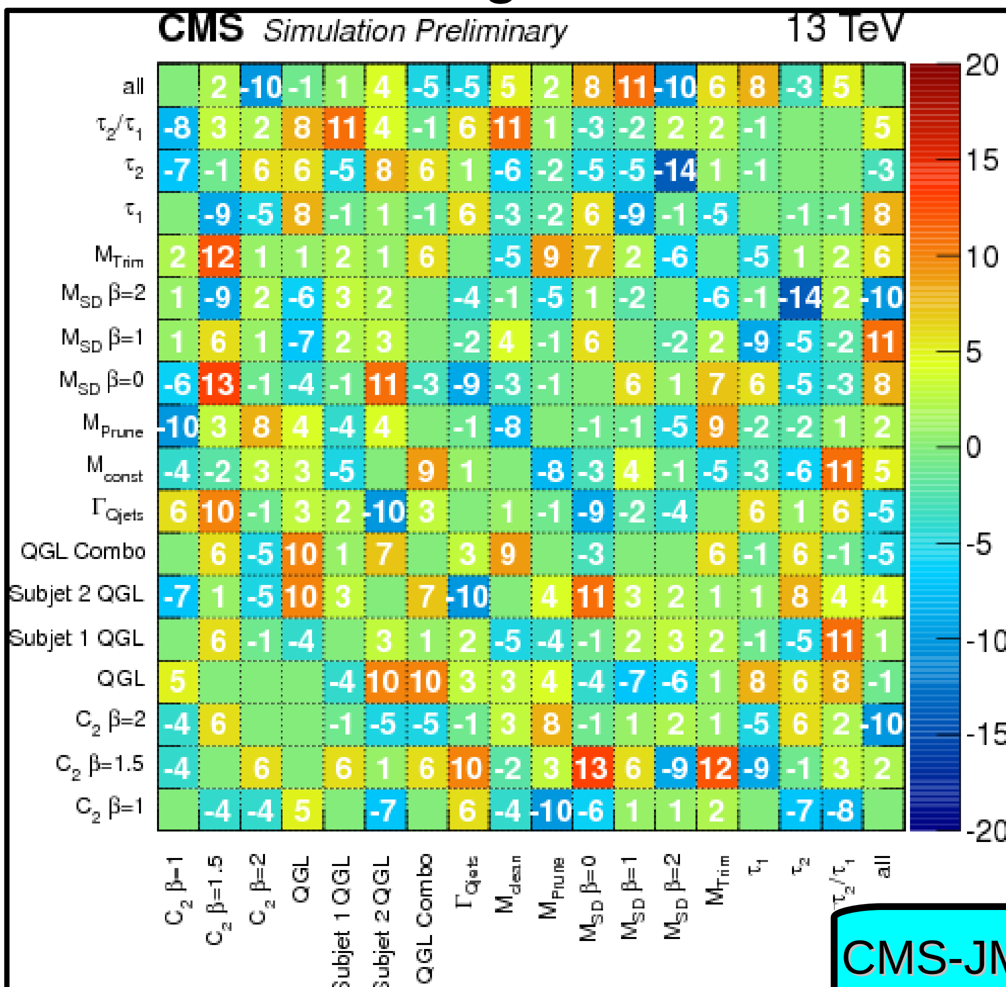
τ_2/τ_1, Qjet, $C_2(\beta = 1.5)$,
Q/G Like on sub-jet 2

Pileup dependence??

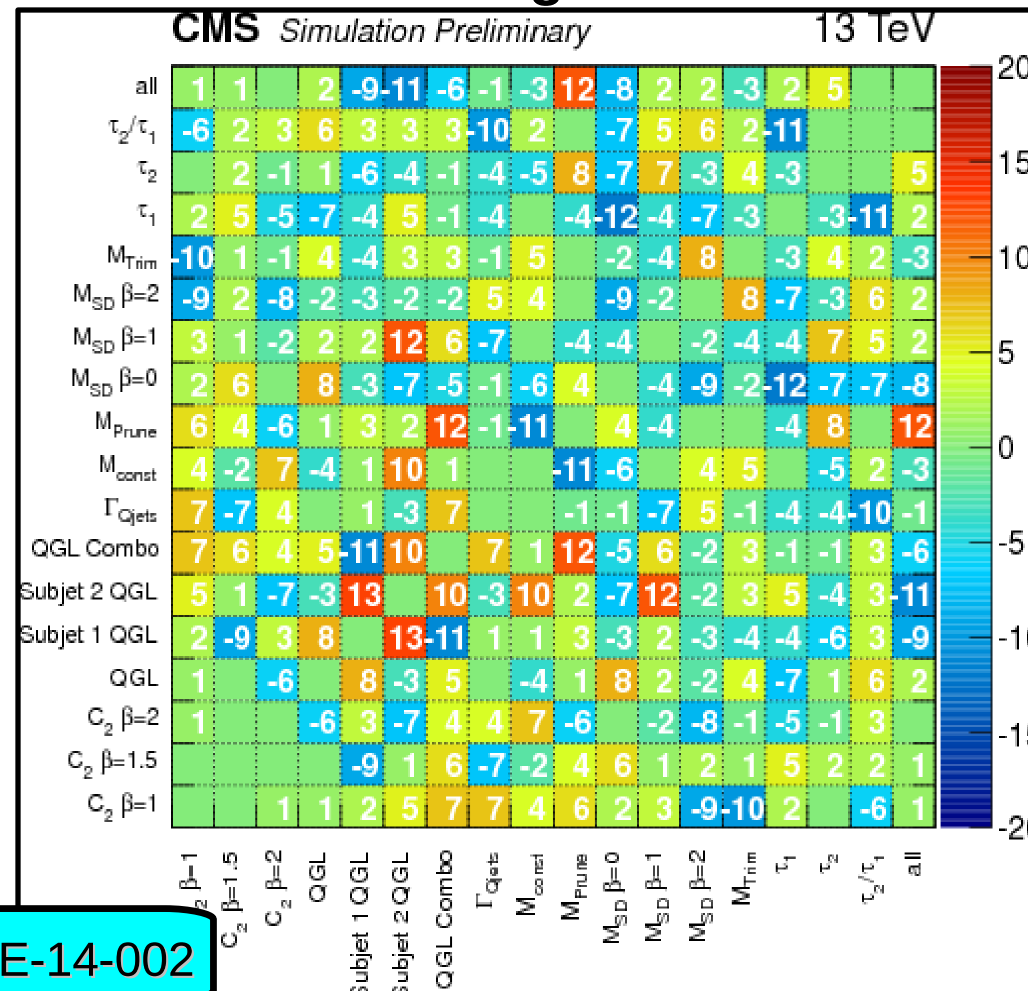
BDT Correlation vs pile-up

- Compute **linear correlation** between single variable BDTs and full combination
- Look at the difference in the correlation between lower and higher pileup bins for S and B

Signal



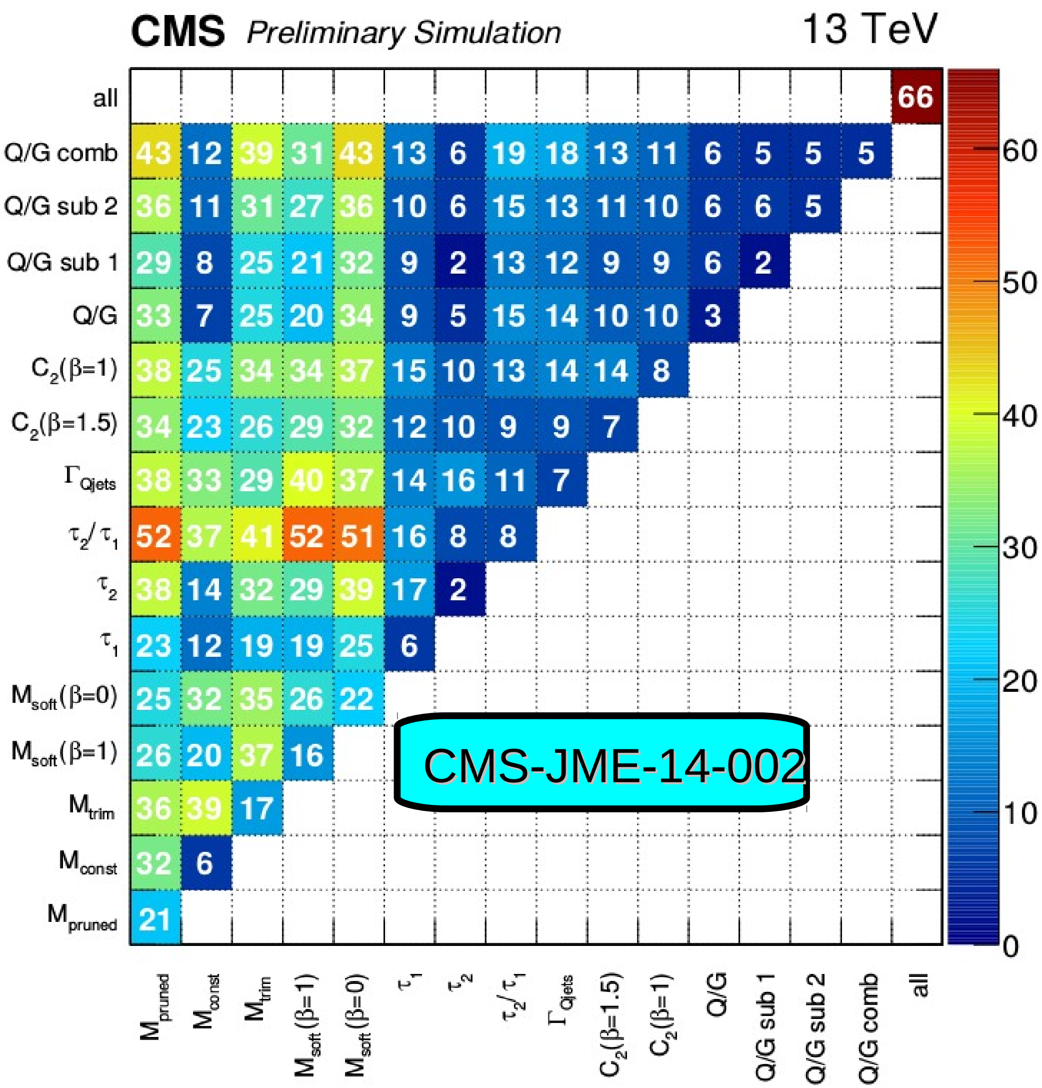
Background



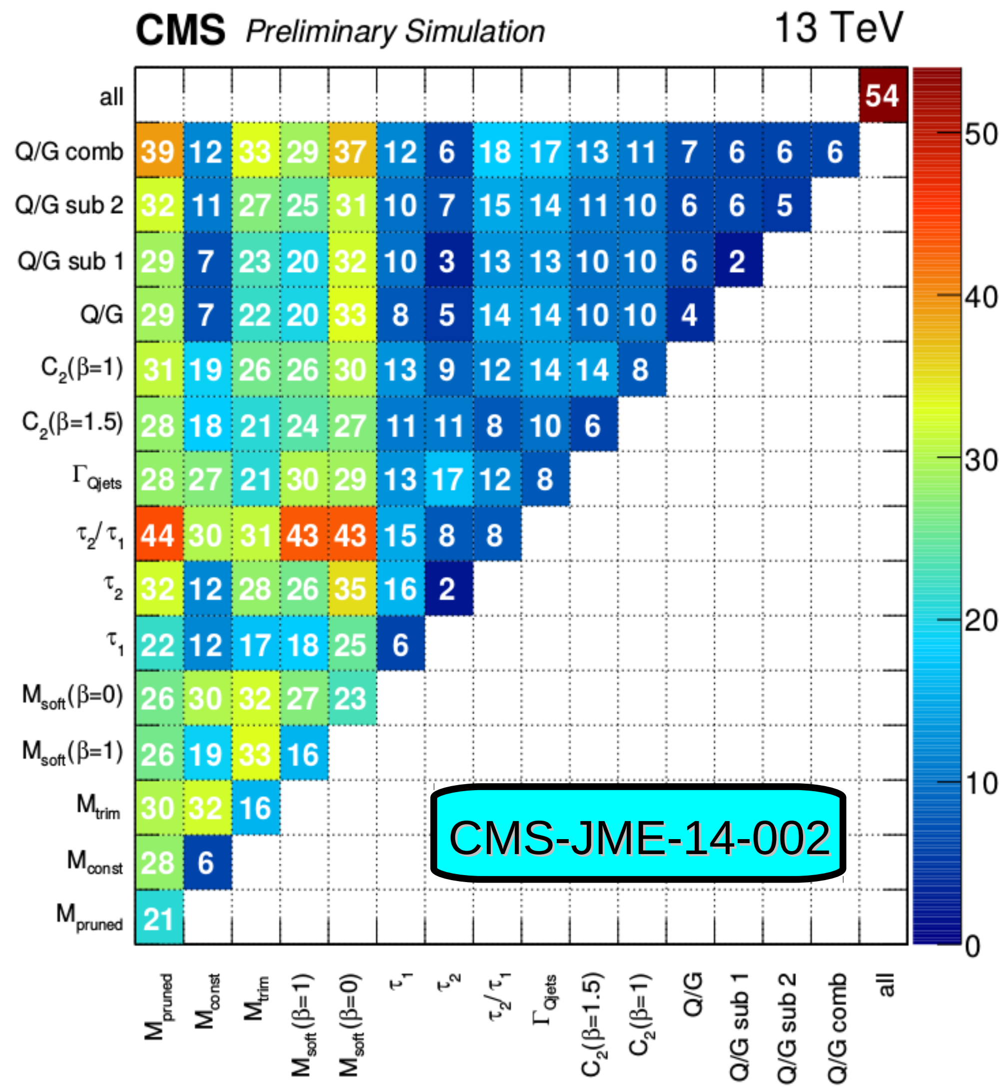
CMS-JME-14-002

BDT performance: pairwise

Low Pileup

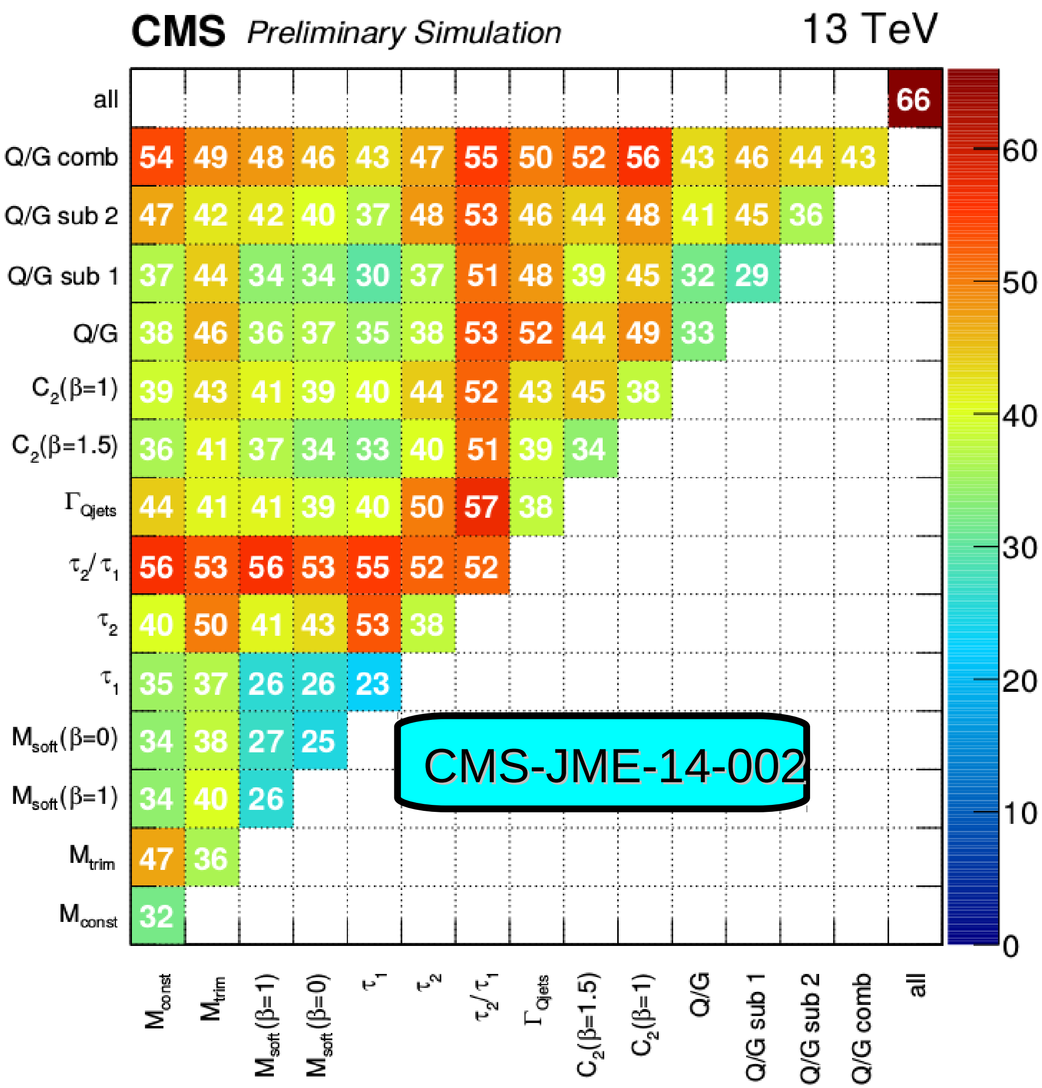


high Pileup

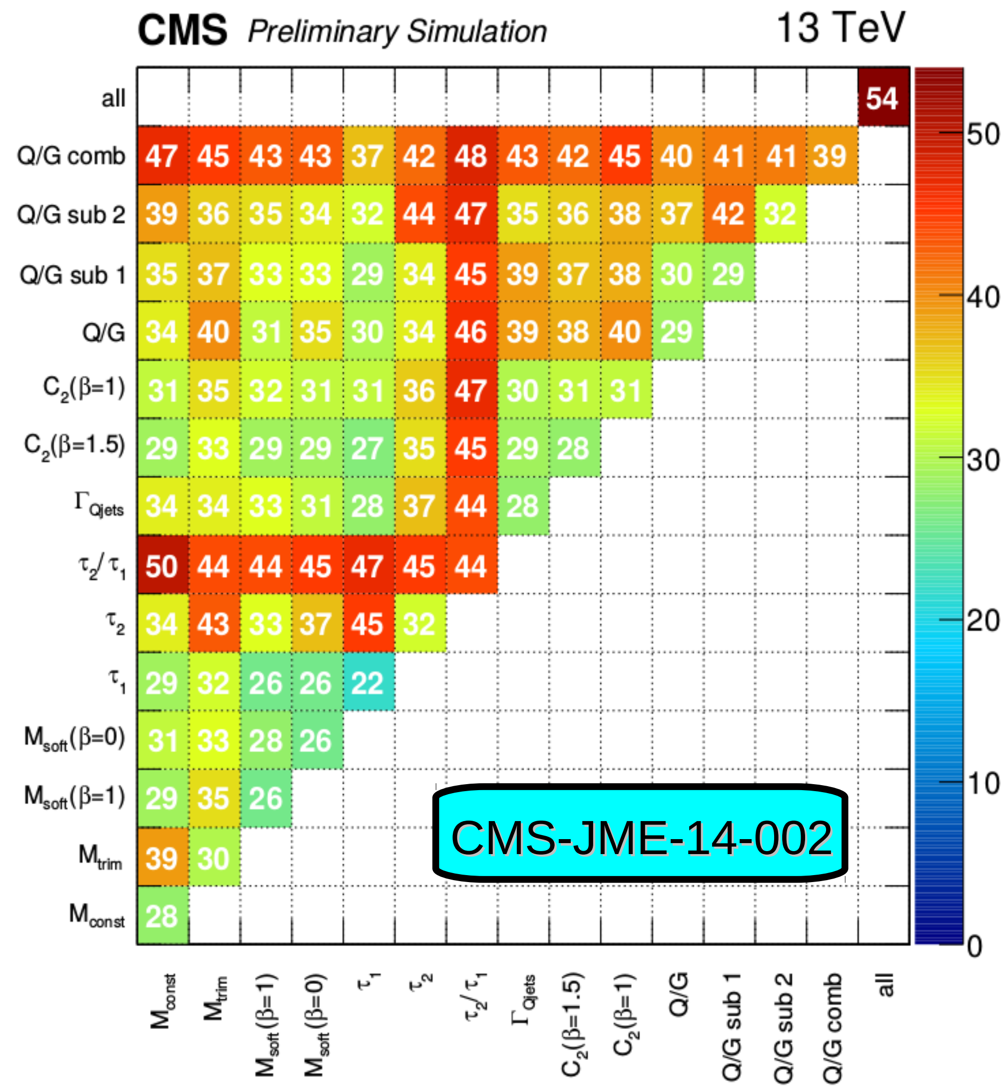


BDT performance: triplets

Low Pileup



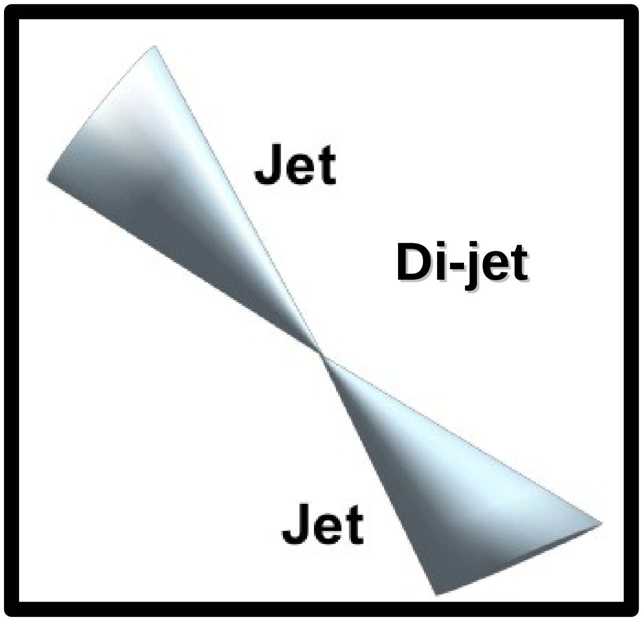
high Pileup

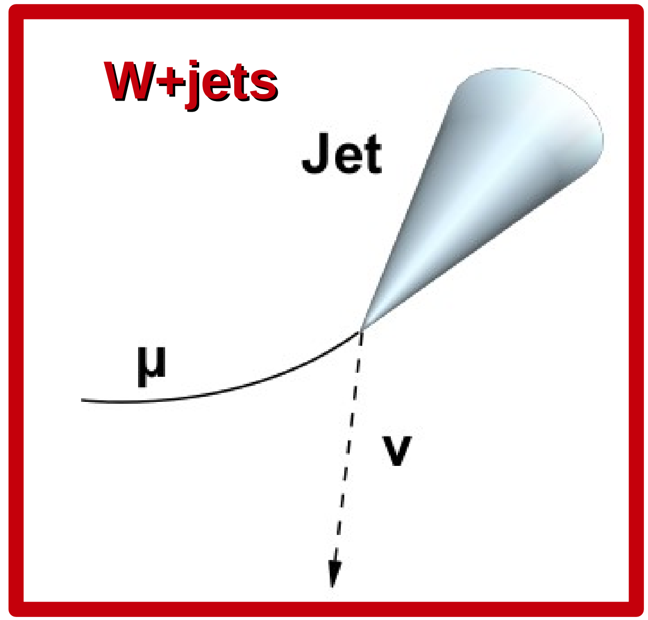


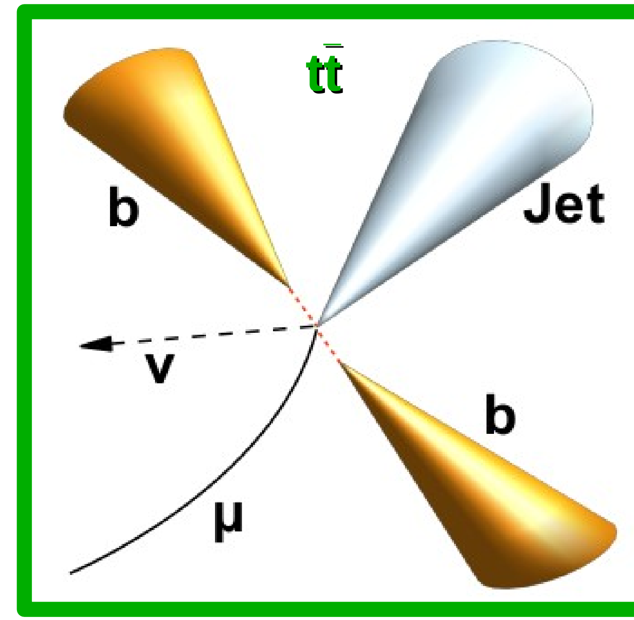
Boosted W/Z performance JME-13-006

Boosted W/Z objects in CMS

So far CMS analysis used CA, R = 0.8 + charged hadron subtraction as “seed” for boosted W/Z decaying hadronically







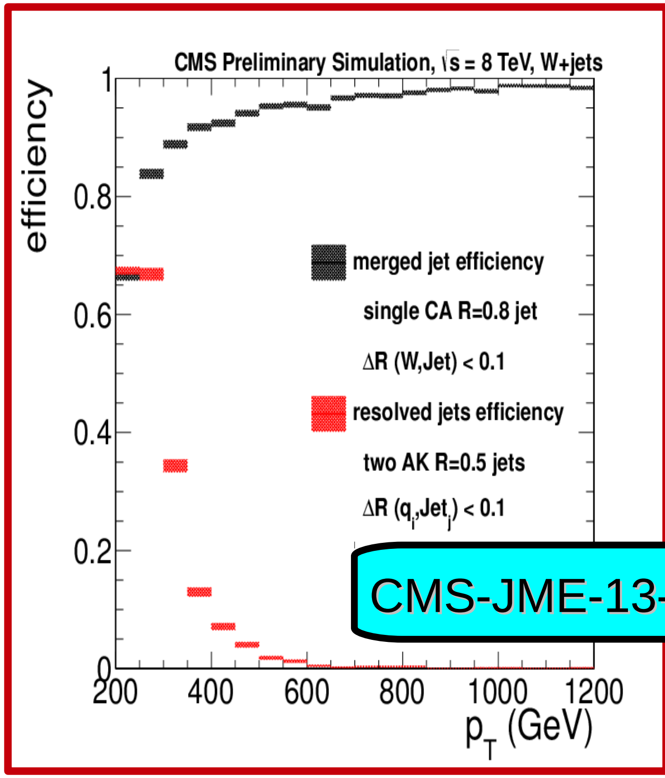
- Two hard jets p_T = [0.4,0.6] TeV
- High production cross section
- q/g background sample

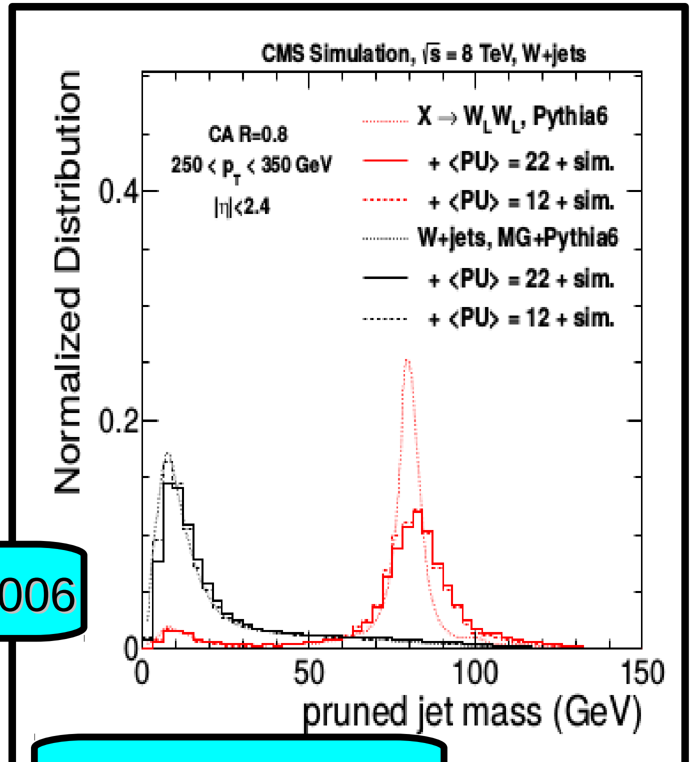
- Low p_T [0.25,0.35] TeV
- QCD jet dominated
- contamination: $t\bar{t}$, VV, DY

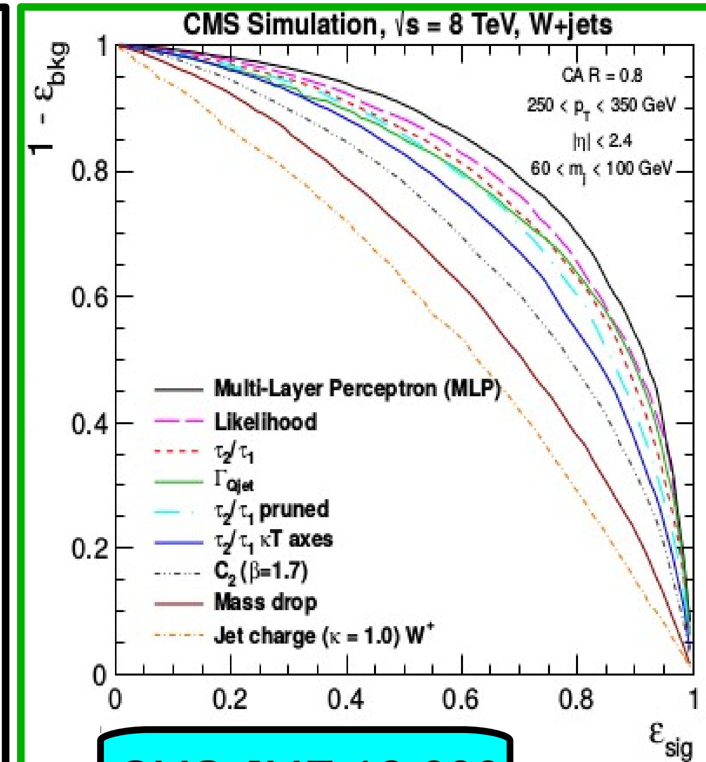
- Isolate W-jet sample in data
- Combinatorial bkg due to b-jet
- Nearby b-jet effect at high p_T

Beyond SM resonances are used as source of W jets

Boosted W/Z in CMS: performance







- Resolved jet from AK5 less efficient from $p_T > 200$ GeV
- Overlap between boosted and resolved searches

- Pruning used improve S/B discrimination

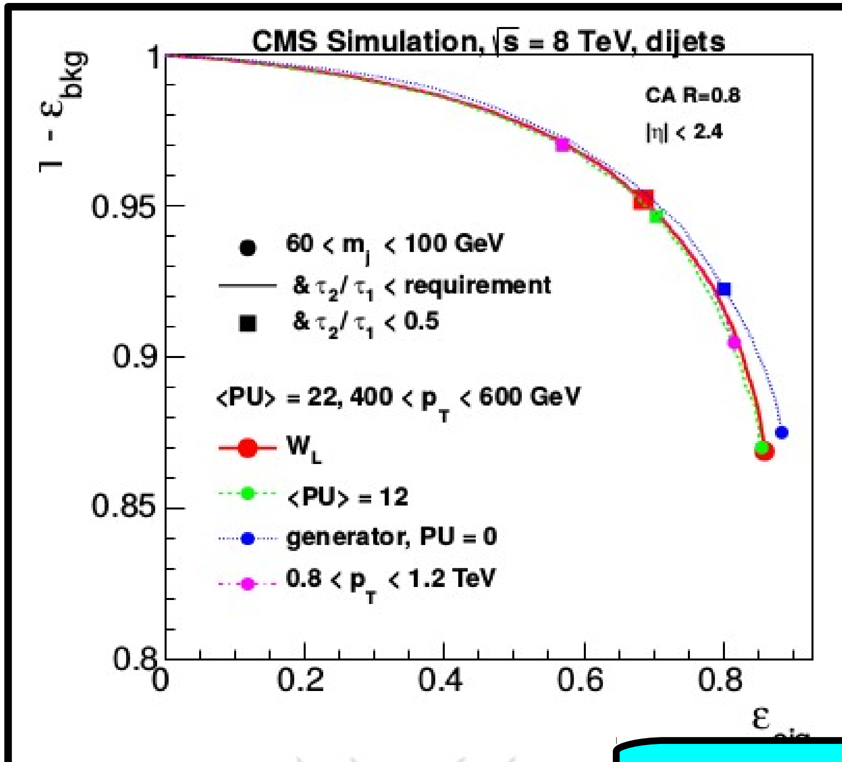
  $Z_{cut} = 0.1$, $R_{cut} = 0.5$

- Best discrimination → **Mass window [60,100] GeV**

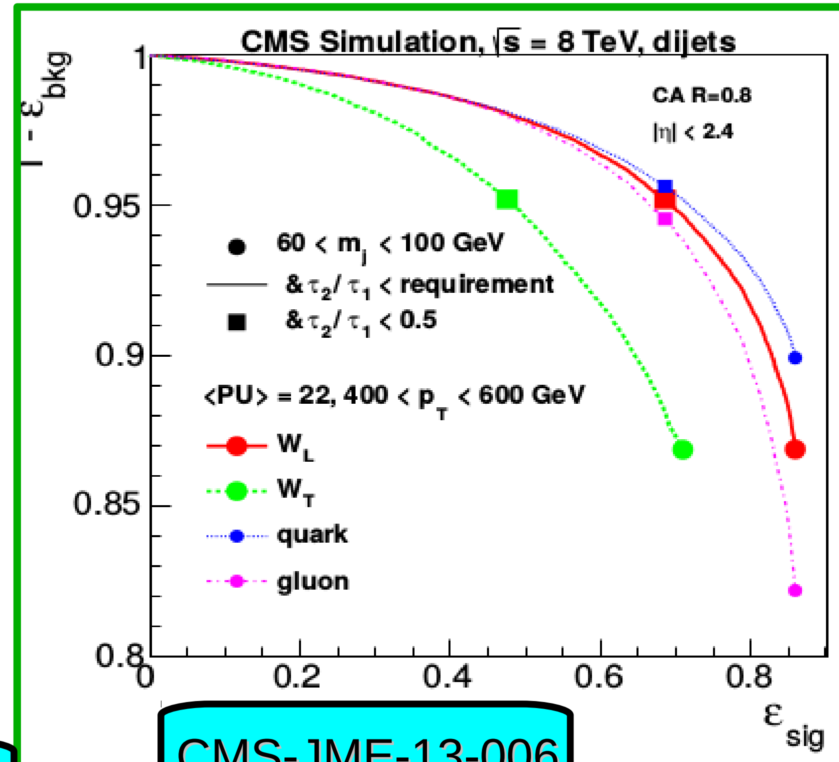
- **Mass window [60,100] GeV**
- Look at the performance of other substructure observables
- **CMS Working Point:**

Mass [60,100] GeV + $\tau_2/\tau_1 < 0.5$

Boosted W/Z in CMS: systematics



CMS-JME-13-006



CMS-JME-13-006

- Higher p_T → same 1 - ε(bkg) moving τ_2/τ_1 cut

  Fixed WP degrades performance

- Worsening due to pileup w.r.t generator level

✔ **W polarization has a big impact**

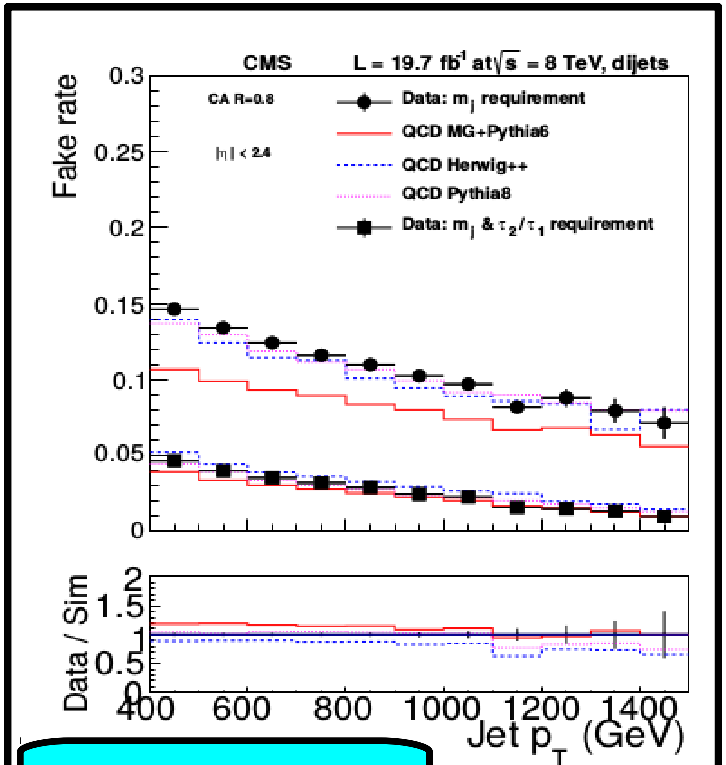
Different angular and p_T sub-jet distributions

✔ **Background composition q/g**

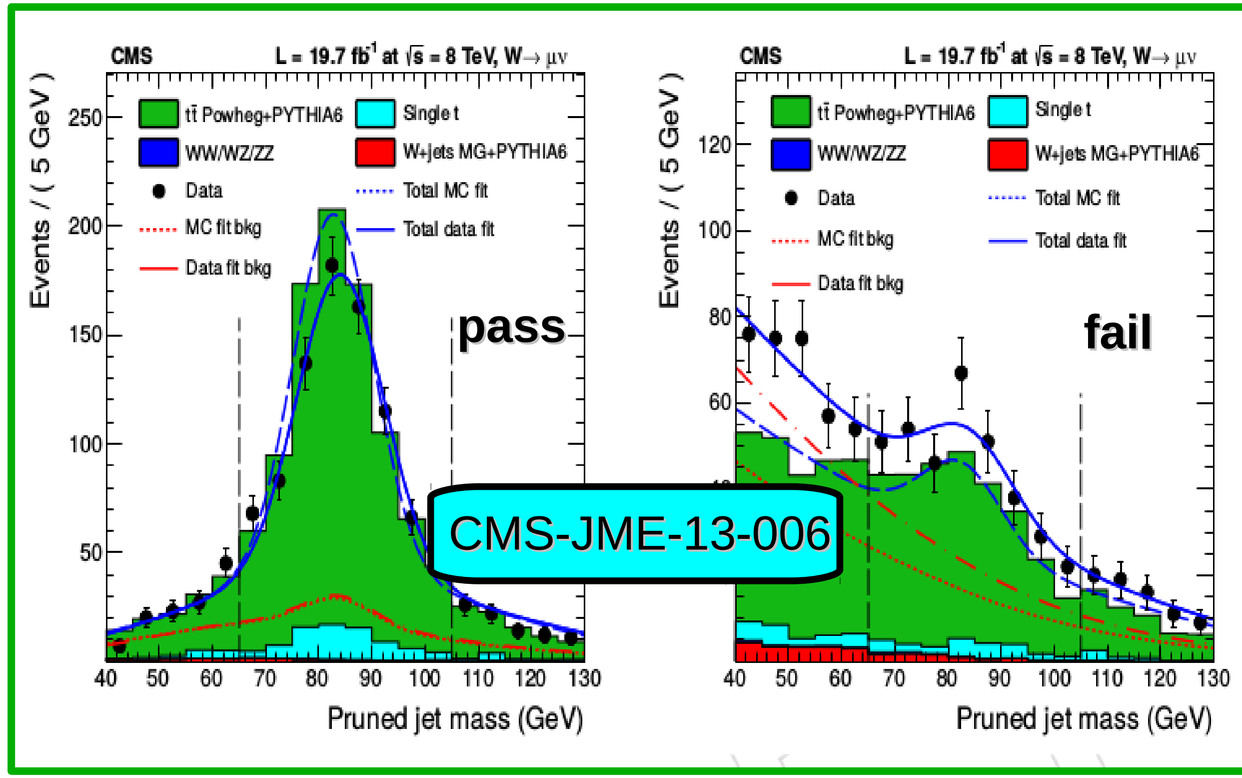
Better rejection for q jets w.r.t gluons

Mass cut reject more quarks ; τ_2/τ_1 gluons

Boosted W/Z in CMS: data analysis



CMS-JME-13-006



- **Mistag rate in data vs p_T**

  Di-jet sample for M_{pr} and M_{pr} + τ_2/τ_1

  Drop vs p_T ; same for signal efficiency

- **Mistag vs PU stable at 1% level**

✔ **Mesurement of W-tagging efficiency in data**

Semi-leptonic $t\bar{t}$ control sample

Simultaneous fit to jet mass for event passing and failing $\tau_2/\tau_1 < 0.5$

Shapes from MC matched/unmatched W-jet with Gen W

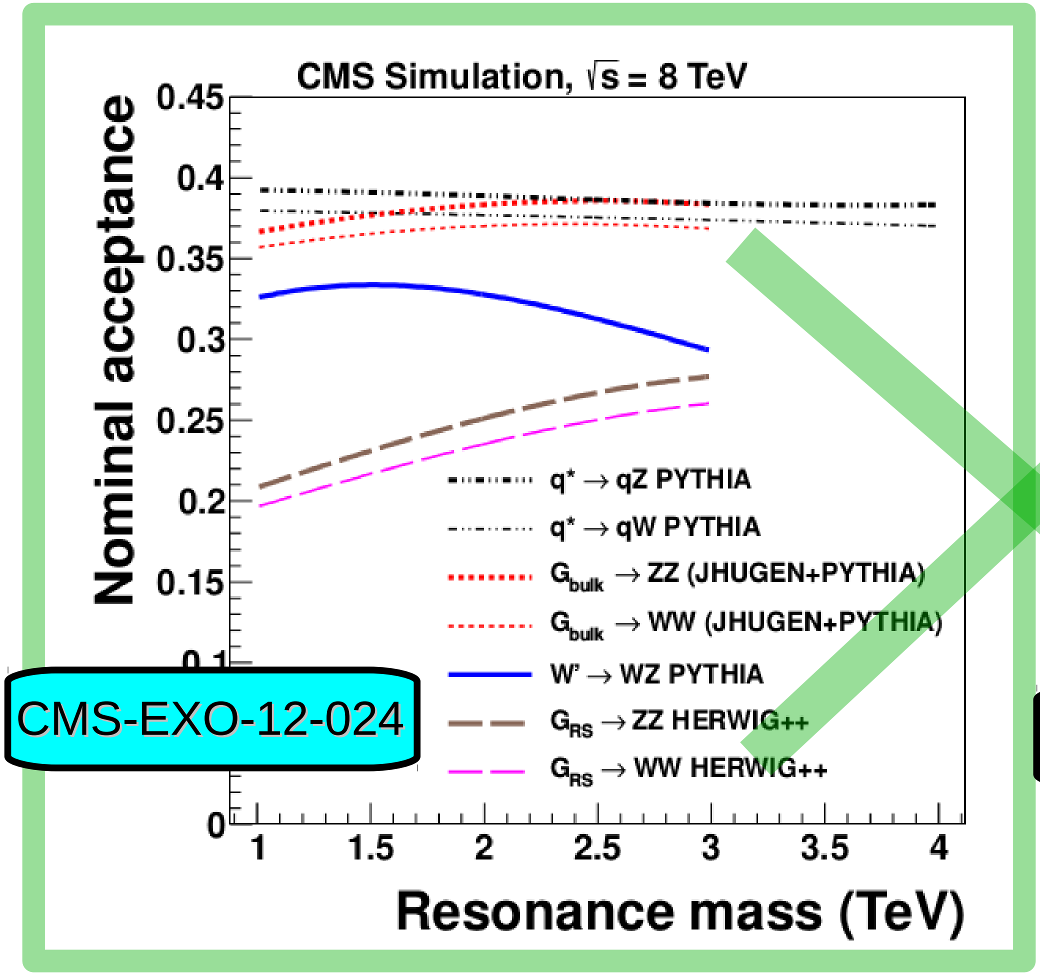
Parameter	Data	Simulation	Data / Simulation
$\langle m \rangle$	84.1 ± 0.4 GeV	82.7 ± 0.3 GeV	1.017 ± 0.006
σ	8.4 ± 0.6 GeV	7.6 ± 0.4 GeV	1.11 ± 0.09

Di-jet search EXO-12-024

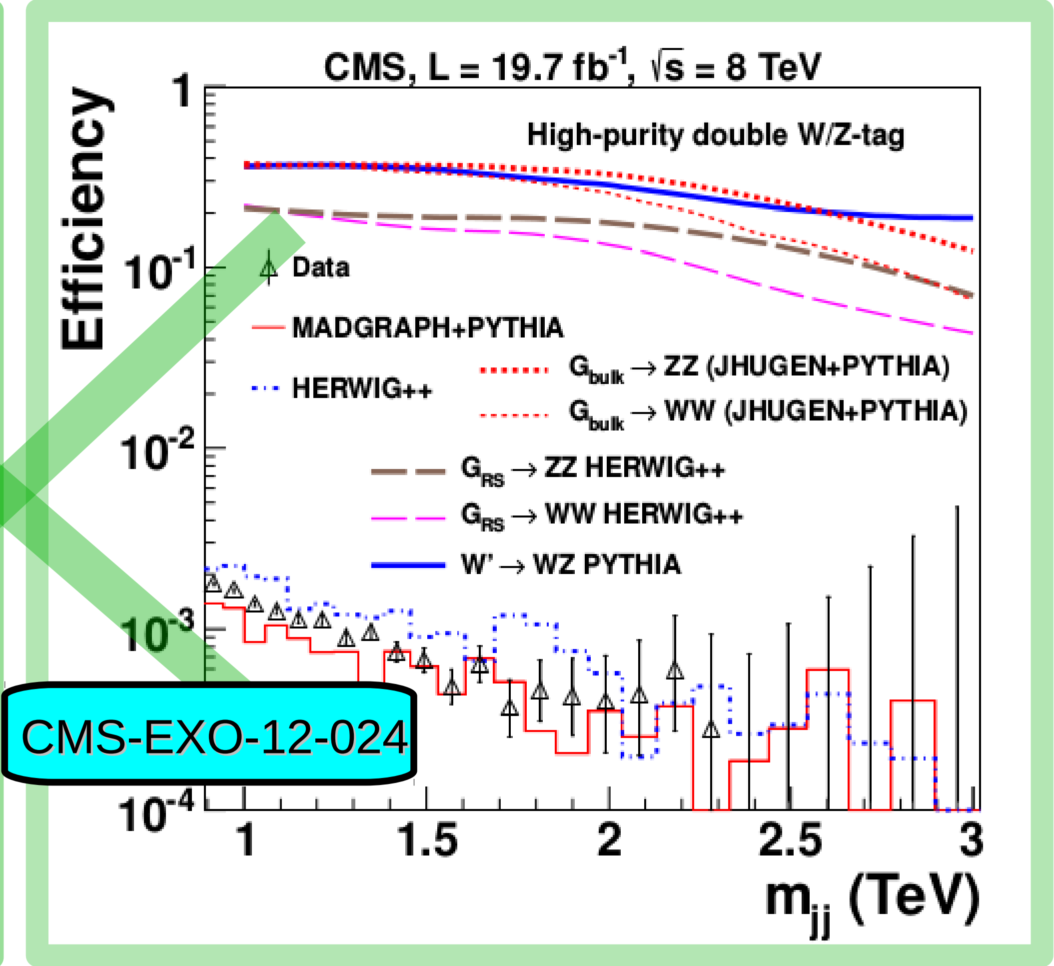
X→VV→di-jet : model independent

- **Goal** → estimate **global event selection efficiency** for signal using full simulated events
- **Contamination** → less than 1% from semi-leptonic WW/WZ and ZZ, 3% from ZZ →qq ττ

Nominal Acceptance → includes matching and jet reconstruction efficiency

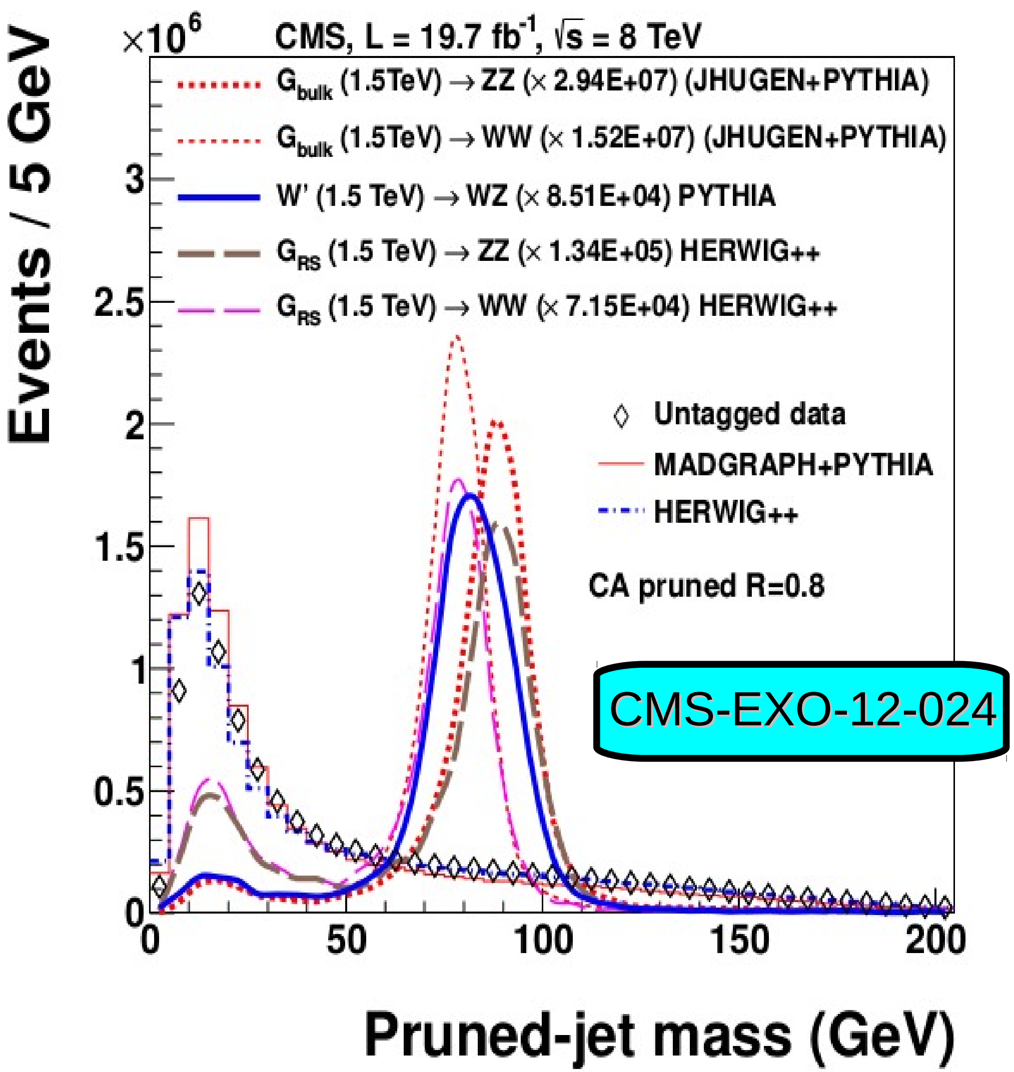


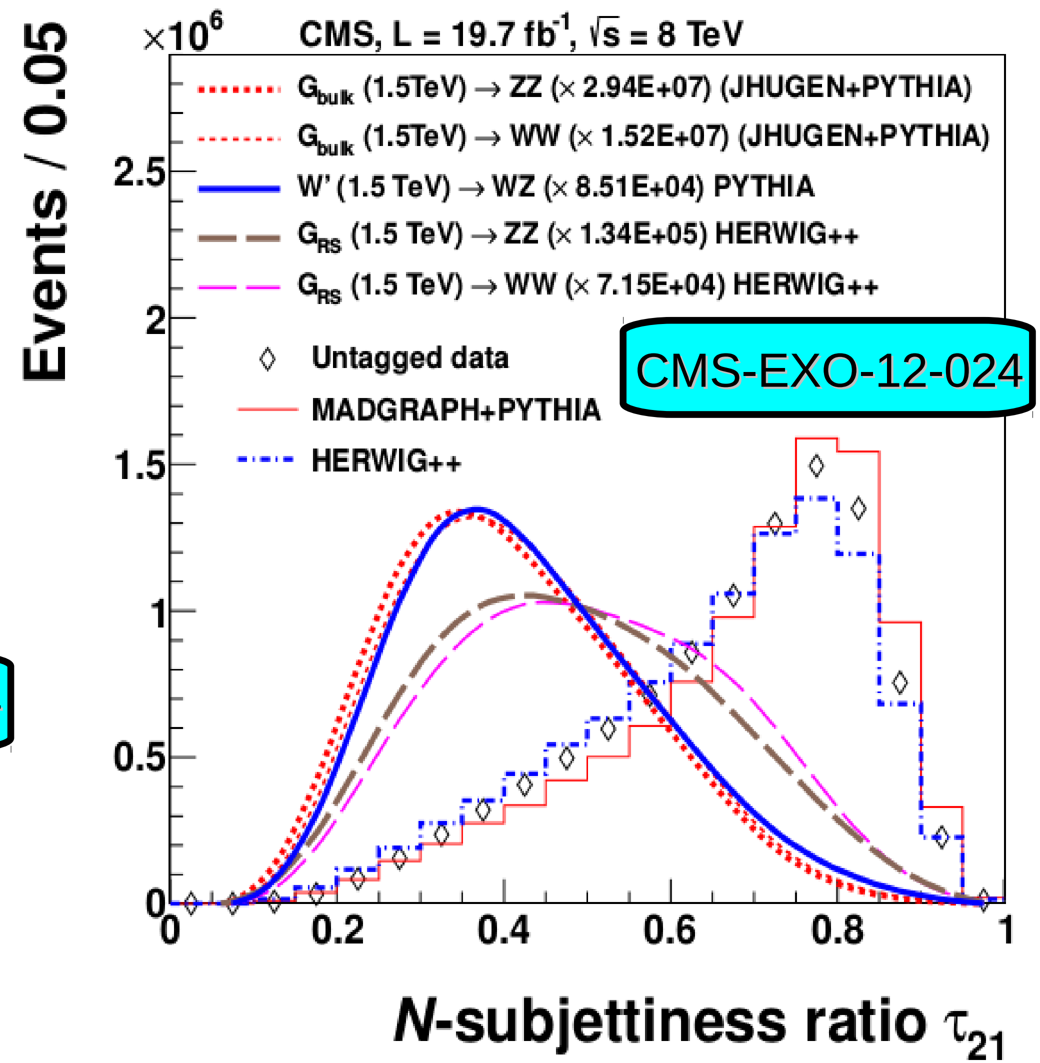
W/Z efficiency → τ_{21} drops at high p_T



X→VV→di-jet search: W/Z tagging

- **Background** → as known Herwig++ agrees better than Pythia6 with data
- **Signal** → different shapes due to **parton shower, W/Z polarization and jet p_T**





X → VV → di-jet : systematic uncertainties

- **Background →** **data driven extraction → one source associated to the method itself**
- **Signal →**
  1) W/Z tagging efficiency scale factor from tag & probe + larger p_T extrapolation.
  2) Jet Energy Scale (JES) and Jet Energy Resolution (JER).
  3) Integrated luminosity.

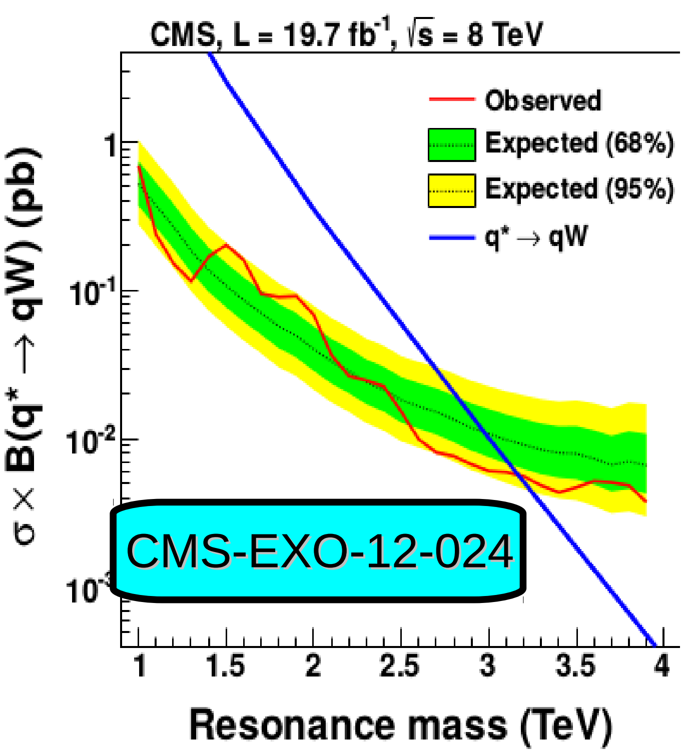
Table 1: Summary of systematic uncertainties. The labels HP and LP refer to high-purity and low-purity event categories, respectively.

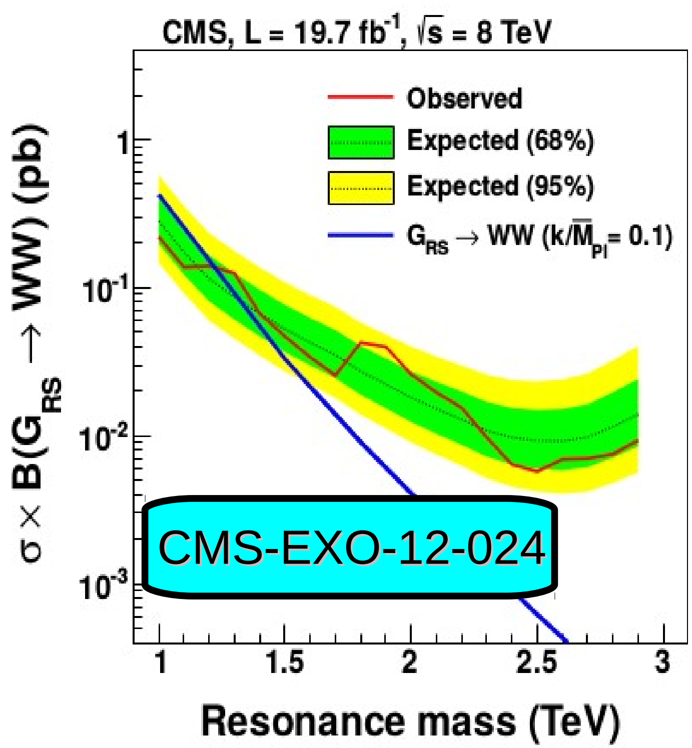
Source	Relevant quantity	LP uncertainty (%)	HP uncertainty (%)
Jet energy scale	Resonance shape	1	1
Jet energy resolution	Resonance shape	10	10
W-tagging	Efficiency (per jet)	7.5	54
Tagging p_T-dependence	Efficiency (per jet)	<4	<12
Pileup	Efficiency (per jet)	<1.5	<1.5
Integrated luminosity	Yield (per event)	2.6	2.6
PDF	Yield (per event)	5–15	5–15

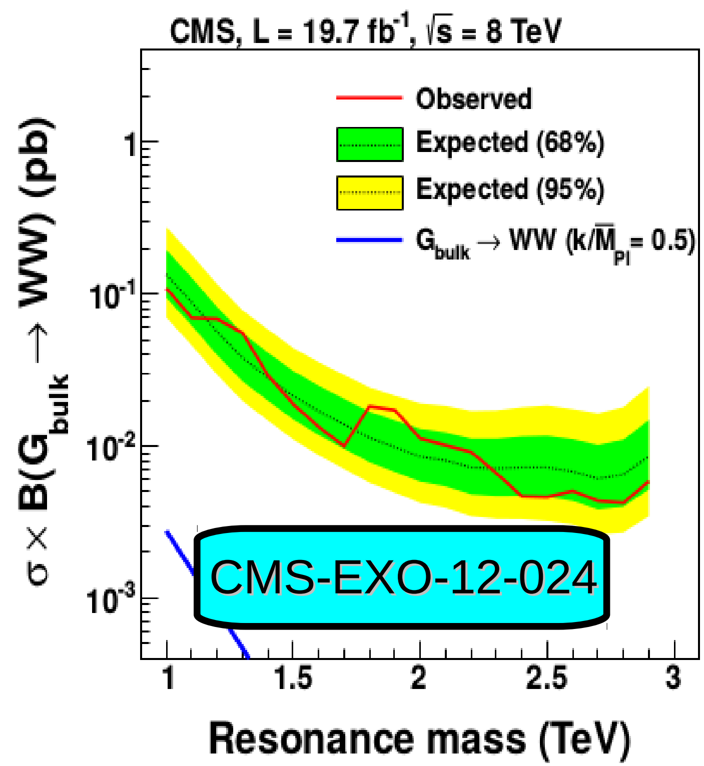
CMS-EXO-12-024

X → VV → di-jet search: results

- HP and LP categories combined in one Likelihood → LHC CLs method to extract upper limit
- **Background uncertainties are uncorrelated ; Signal ones are fully correlated**







Results:

Excited quarks decaying in qW , qZ are excluded till 3.2 and 2.9 TeV, respectively

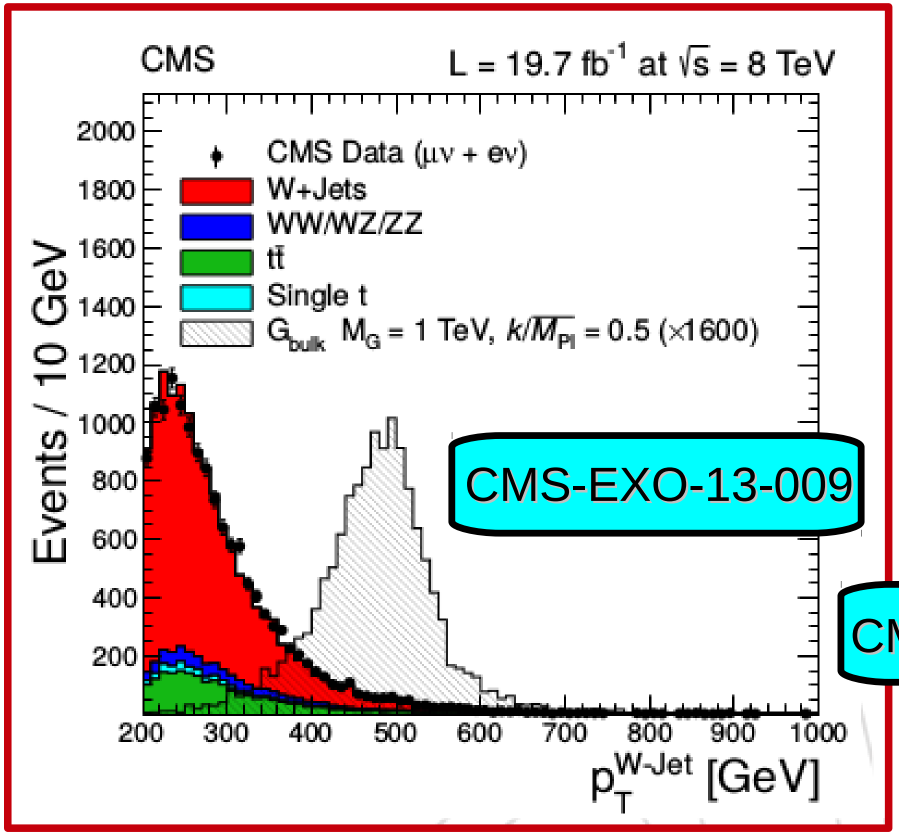
G_{RS} → WW excluded up to 1.2 TeV   W' → WZ up to 1.7 TeV

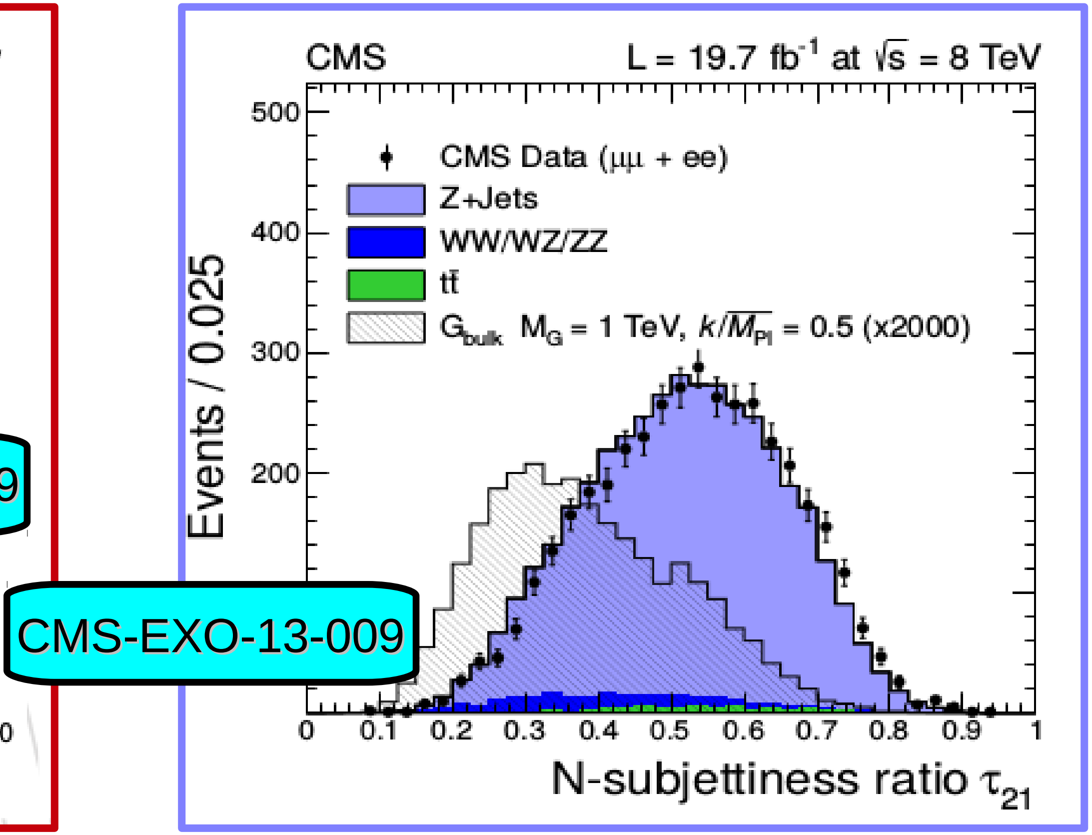
G_{bulk} → WW, WZ, ZZ no mass limit up to now

Lepton+jet search EXO-13-009

$X \rightarrow VV \rightarrow$ semi-leptonic : event selection

- **l + ν + jet** → leptonic W mass constraint used to extract the neutrino p_Z^{ν}
  Missing transverse energy forced to be a measurement of p_T^{ν}
  $p_T > 50$ (90) GeV μ (e) ; MET > 40 (80) GeV μ (e) ; $p_T^W > 200$ GeV
- **2l + jet** → $p_T > 40$ (20) GeV for leading (trailing) μ ; $p_T > 40$ GeV ele ; $p_T^W > 100$ GeV
  Two opposite charge lepton, M_{ll} in [70,110] GeV → **kill top background**
- **Hadronic W identification** → same strategy adopted in di-jet searches





X → VV → semi-leptonic : systematics

- **Background → normalization:** 1) V+jets from limited statistics in the SB region **< 10%**
  2) $t\bar{t}$ from correction derived from control region ~ **5-7%**
  3) VV cross section uncertainty assigned to be **20%**
  4) V-tagging scale factor **~10% for HP category**

  → **V+jet shape:** covariance matrix of the fits done in SB region and $\alpha_{MC}(M_{VV})$

- **Signal → shape:** 1) Jet Energy Scale and Resolution effect on signal width 3% (2%)
  2) Lepton scale and resolution give small effect on peak and width < 1%

  → **normalization:**

  1) physics object uncertainties are assumed to be uncorrelated
  2) Trigger and ID systematics from dedicated tag&probe studies (Z→ll)
  3) Luminosity from CMS measurement
  4) PDF: MSTW and NNPDF

Source	Analysis $\ell\nu$+V-jet	Analysis $\ell\ell$+V-jet
Muons (trigger and ID)	2%	5%
Muon scale	1%	2%
Muon resolution	<0.1%	0.5%
Electrons (trigger and ID)	3%	3%
Electron scale	<0.5%	<0.5%
Electron resolution	<0.1%	<0.1%
Jet scale	1–3%	1%
Jet resolution	<0.5%	<0.1%
Unclustered energy scale	<0.5%	—
Pileup	0.5%	0.5%
V tagging	9% (HP) 24% (LP)	
PDF	<0.5%	
Luminosity	2.6%	

CMS-EXO-13-009

X → VV → semi-leptonic : M_{VV} shape

V+jets shape in the signal region is extracted from data in the low jet mass sideband

Unbinned fit of data in the M_{VV} SB with a falling shape

Extrapolation function from simulation used to obtain the shape into the signal region

$$\alpha_{MC}(m_{VV}) = \frac{F^{V+jets}_{MC,SR}(m_{VV})}{F^{V+jets}_{MC,SB}(m_{VV})},$$

High mass SB not used to exclude contamination from possible VH signals

Final background estimation in the signal region

$$N^{BKGD}_{SR}(m_{VV}) = C^{V+jets}_{SR} \times F^{V+jets}_{DATA,SB}(m_{VV}) \times \alpha_{MC}(m_{VV}) + \sum_k C^k_{SR} F^k_{MC,SR}(m_{VV}).$$

$$e^{-x/(c_0+c_1 x)}$$

N_{SR}^{BKG} → total background in SR as function of M_{VV}

C_{SR}^{V+jets} → V+jets normalization from m_J fit to data

C_{SR}^{k} → yields of minor background in SR

Fitting Range:

1) l+ν+jet M_{VV} [700,3000] GeV

2) l+l+jet M_{VV} [500,2800] GeV

$X \to VV \to$ semi-leptonic : acceptance

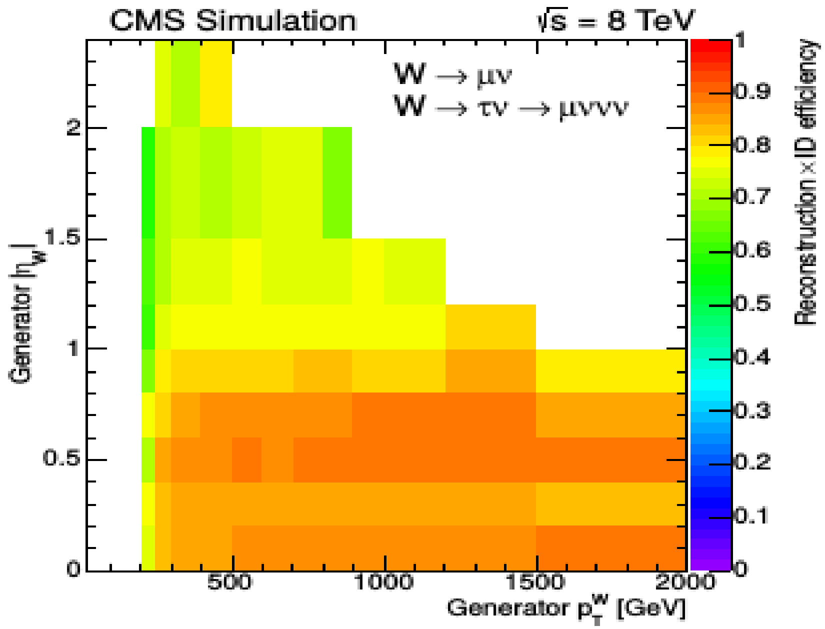
l+ν+jet acceptance cuts

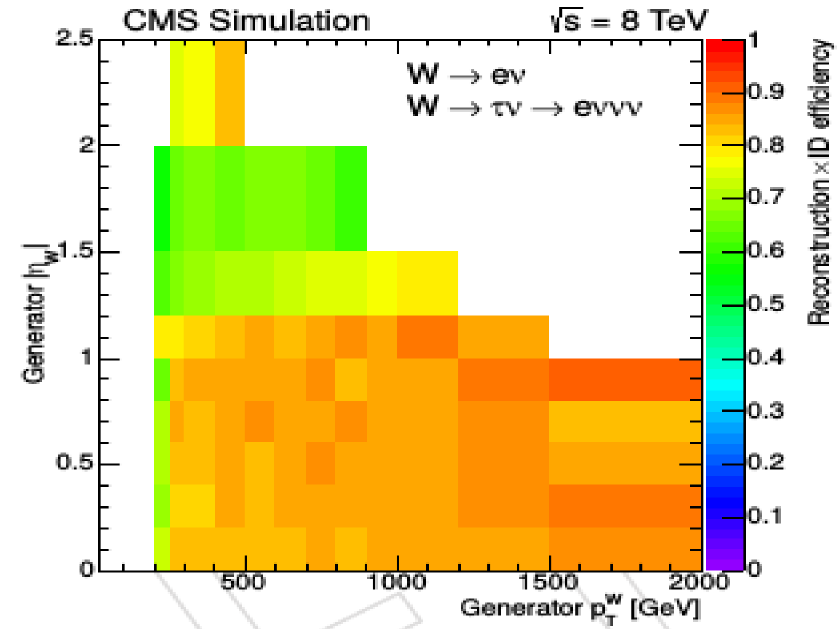
Object	Requirement
Muons	$\lvert\eta\rvert < 2.1$
	$p_T > 50\,GeV$
Electrons	$\lvert\eta\rvert < 2.5$
	$p_T > 90\,GeV$
$\sum \vec{p}_{T,\nu}$ (Muon ch.)	$p_T > 40\,GeV$
$\sum \vec{p}_{T,\nu}$ (Electron ch.)	$p_T > 80\,GeV$
$W \to \ell\nu$ or $W \to \tau\nu \to \ell\nu\nu\nu$	$p_T^W > 200\,GeV$
$W \to q\bar{q}'$	$\lvert\eta_W\rvert < 2.4$
	$p_T^W > 200\,GeV$
	$65 < m_{q\bar{q}'} < 105\,GeV$
WW system	$700 < m_{WW} < 3000\,GeV$
	$\Delta R(W_{q\bar{q}'}, \ell) > \pi/2$
	$\Delta\phi(W_{q\bar{q}'}, \sum \vec{p}_{T,\nu}) > 2$
	$\Delta\phi(W_{q\bar{q}'}, W_{\ell\nu}) > 2$

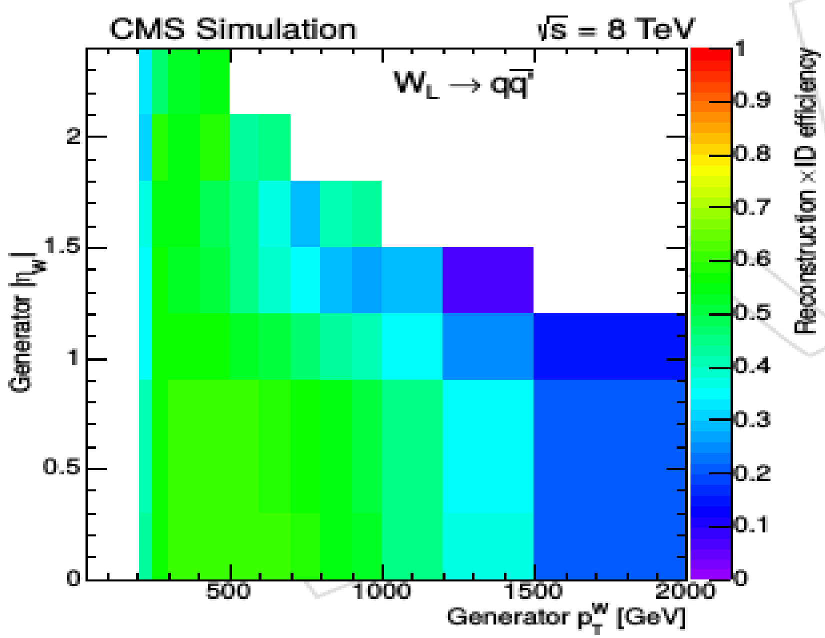
l+l+jet acceptance cuts

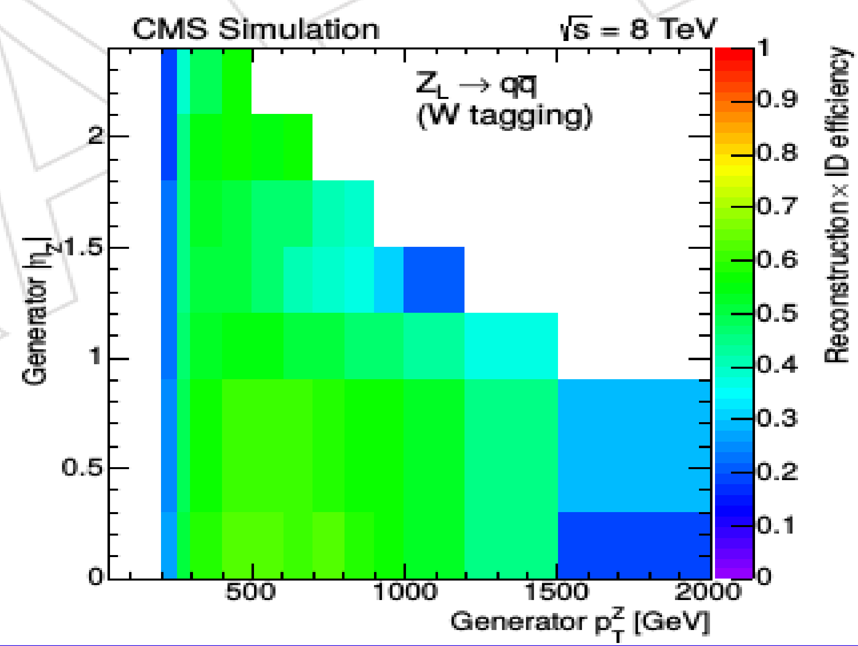
Object	Requirement
Muons	$\lvert\eta\rvert < 2.4$
	$p_T > 20\,GeV$
Highest-p_T muon	$p_T > 40\,GeV$
Electrons	$\lvert\eta\rvert < 2.5$
	$p_T > 40\,GeV$
$Z \to \ell\ell$	$p_T^Z > 80\,GeV$
	$70 < m_{\ell\ell} < 110\,GeV$
$Z \to q\bar{q}$	$\lvert\eta_Z\rvert < 2.4$
	$p_T^Z > 80\,GeV$
	$70 < m_{q\bar{q}} < 110\,GeV$
ZZ system	$500 < m_{ZZ} < 2800\,GeV$

X → VV → semi-leptonic : efficiency WW

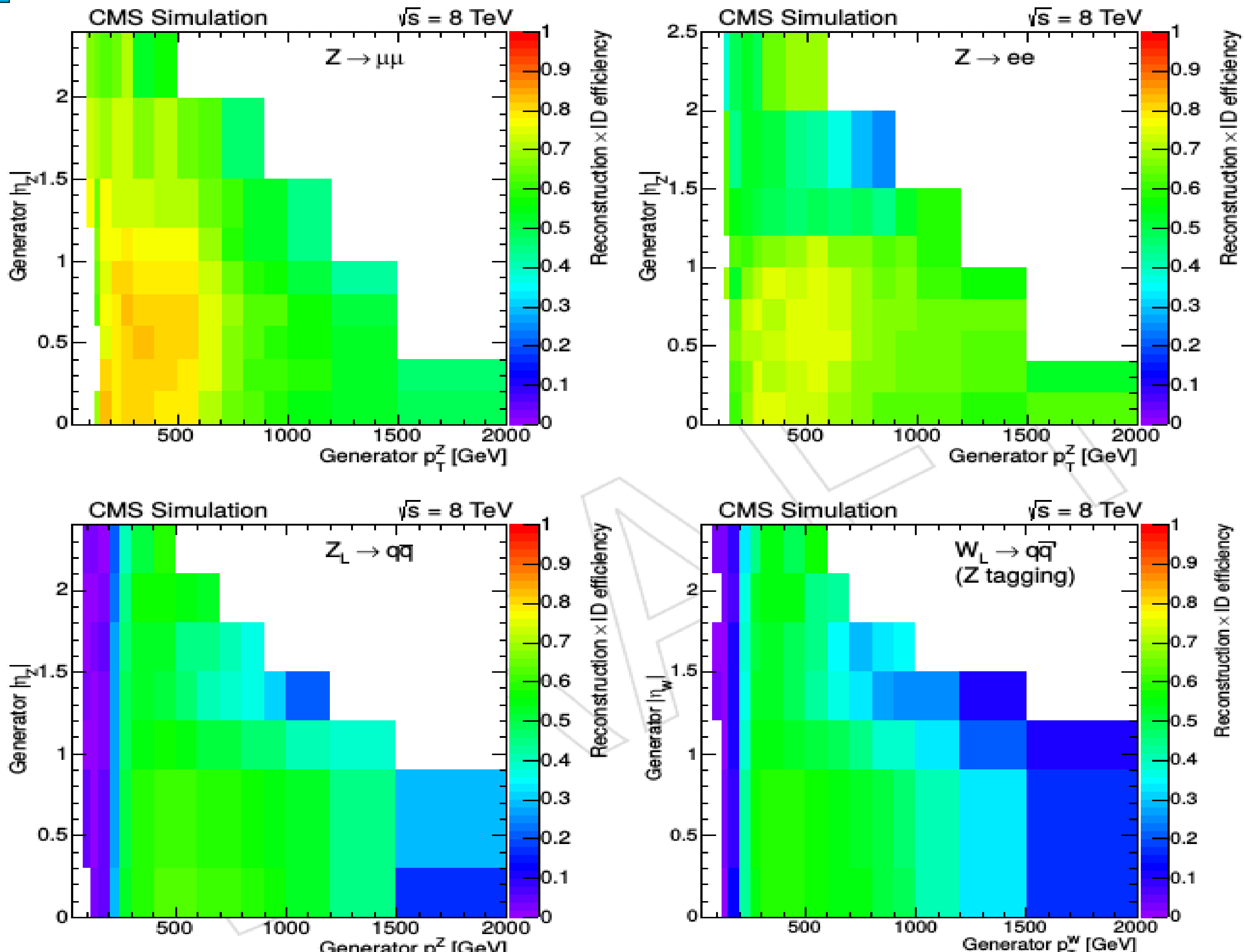








X → VV → semi-leptonic : efficiency ZZ



Additional B2G searches

T(2/3) quark search

Signal:

Simulated by **MG+Pythia6**

Model:
exotic T(2/3) coupled with gluons

NNLO cross section calculated via ***HATHOR***

Signature:

T(2/3) → bW or tZ or tH → no specific BR are assumed

$T\bar{T}$ → single lepton or multi-lepton final states

Offline selection:

CA8 used to look for top-jets, p_T > 200 GeV passing CMS combined Top Tagger
Important for T → tZ and tH events

For hadronic W identification, CA8 p_T > 200 GeV pruned mass [65,105] GeV + N-subjettines $\tau_2/\tau_1 < 0.5$ →
Important for T → bW events

Event classified in **Single lepton** and **Multi-Lepton** channels

Backgrounds

SM W+jets, Z+jets, $t\bar{t}W$, $t\bar{t}Z$, $t\bar{t}$, single-top, WW, WZ, ZZ and $t\bar{t}H$

T(2/3) quark search: single lepton

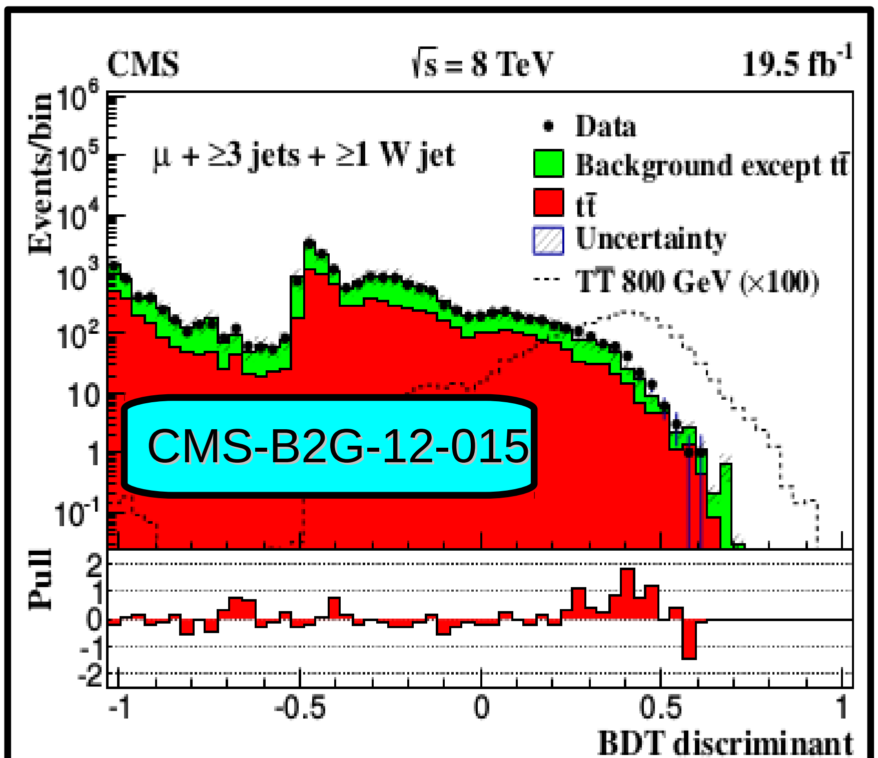
- Lepton p_T > 32 GeV for both muon and electron final state, 2nd lepton veto
- At least three jets with p_T > 120, 90 and 50 GeV ; $\Delta R(j,l) > 0.3$ and $|\eta| < 2.4$
- At least one W-jet with p_T > 35 GeV ; E_T^{miss} > 20 GeV

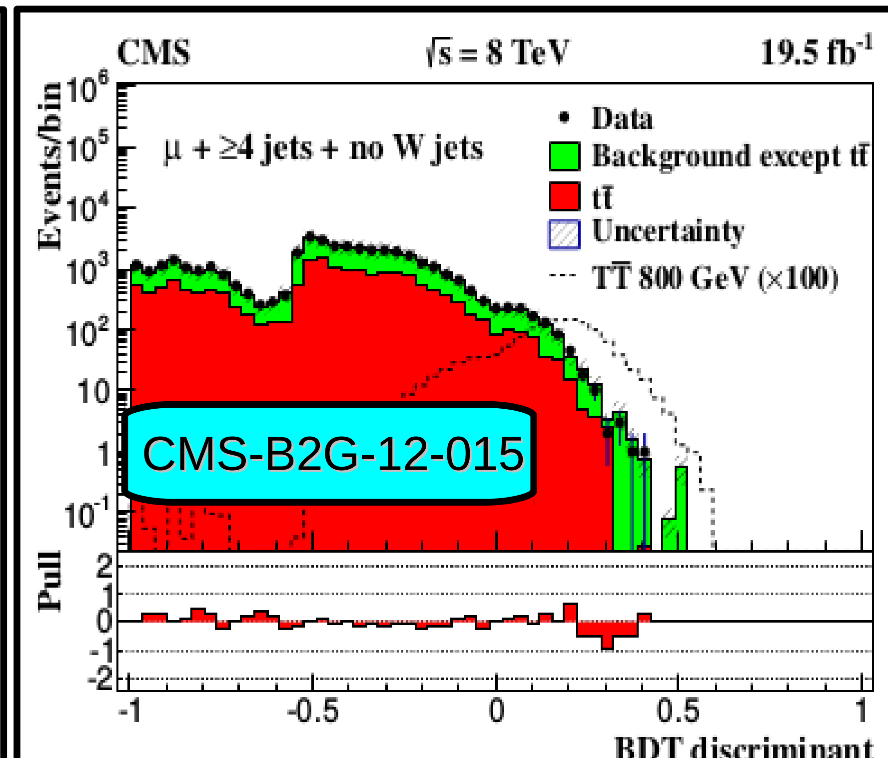
W+jets and $t\bar{t}$ background normalized to data inverting the W-jet requirements

- **BDT analysis to separate T-quark from SM backgrounds ($t\bar{t}$, W and Z+jets)**

One dedicated training for each M_T and for events with or without W-jet

Inputs: N_{jet}, N_{bjet}, H_T, E_T^{miss}, lepton p_T, p_T^{3rd}, p_T^{4th}





T(2/3) quark search: multi lepton

- Divided into 4 exclusive samples asking p_T(lep)> 20 GeV, M_{ll} > 20 GeV E_T^{miss} > 20 GeV
  → **lepton charge** : opposite sign, same sign + **at least one b-jet**
  → **opposite sign** is divided according to the **number of jets** in the event

OS1 → *dominated by $T\bar{T} \rightarrow bWbW$: M_Z veto, M_{lb} > 170 GeV, N_{jet} = 2 or 3, H_T > 300 GeV*

OS2 → *N_{jet} > 5, 2 b-tagged jet, H_T > 500 GeV and S_T > 1 TeV → sensitive to events with a Z*

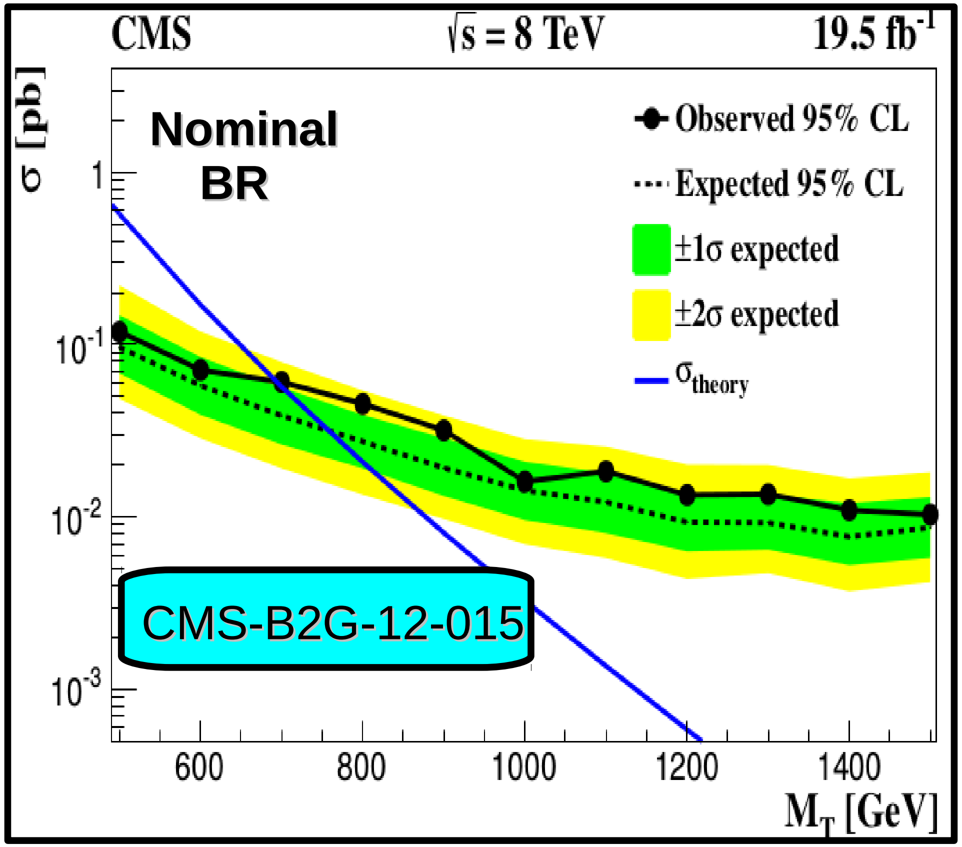
SS → *events with T in tZ or tH, N_{jet} > 3, H_T > 500 GeV, S_T > 700 GeV*

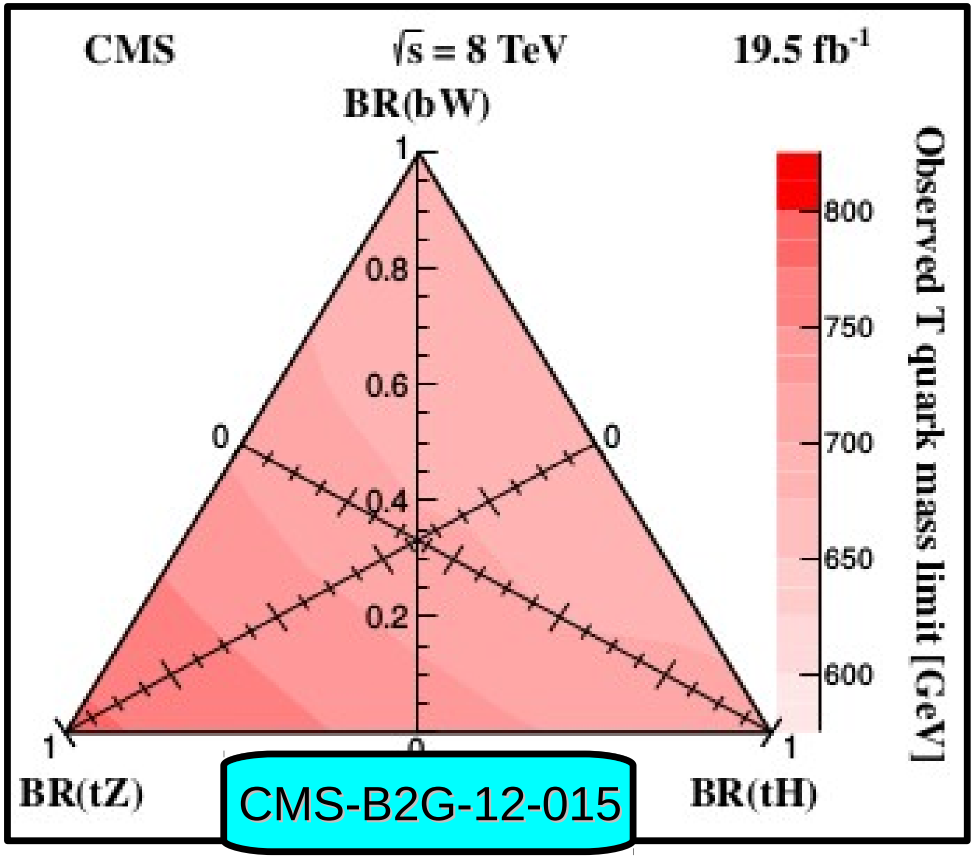
1) SM process with same-sign leptons have small XS → taken from MC
2) Wrong charge measurement → muon negligible, electron fake rate weight method
3) Fake lepton background from dedicated data control sample

Channel	OS1	OS2	SS	Trileptons
$t\bar{t}$	5.2 ± 1.9	80 ± 12	–	–
Single top quark	2.5 ± 1.3	2.0 ± 1.0	–	–
Z	9.7 ± 2.9	2.5 ± 1.9	–	–
$t\bar{t}W$	–	–	5.8 ± 1.9	0.25 ± 0.11
$t\bar{t}Z$	–	–	1.83 ± 0.93	1.84 ± 0.94
WW	–	–	0.53 ± 0.29	–
WZ	–	–	0.34 ± 0.08	0.40 ± 0.21
ZZ	–	–	0.03 ± 0.00	0.07 ± 0.01
WWW/WWZ/ZZZ/WZZ	–	–	0.13 ± 0.07	0.08 ± 0.04
$t\bar{t}WW$	–	–	–	0.05 ± 0.03
Charge misidentification	–	–	0.01 ± 0.00	–
Non-prompt	–	–	7.9 ± 4.3	0.99 ± 0.90
Total background	17.4 ± 3.7	84 ± 12	16.5 ± 4.8	3.7 ± 1.3
Data	20	86	18	2

T(2/3) quark search: results

- Single lepton uses BDT discriminant observed in data to build the bayesian posterior
- Multi-lepton use a cut and count approach





Results:

Mass limits between 687 and 782 GeV as a function of all the BR combination

Analysis more sensitive to tZ mode

T(5/3) quark search

Signals:

Simulated by **MG+Pythia6**

Model:
exotic T(5/3) not coupled with gluons

Signature:

$T\bar{T} \rightarrow 2t + 2W \rightarrow 2W+2b+2W$

Di-lepton (e,μ) same-sign final state

Offline selection:

CA8 used to look for top-jets, $p_T > 400$ GeV passing **CMS combined top tagger**

For hadronic W identification, CA8 $p_T > 200$ GeV pruned mass [65,105] GeV + N-subjettines $\tau_2/\tau_1 < 0.5$

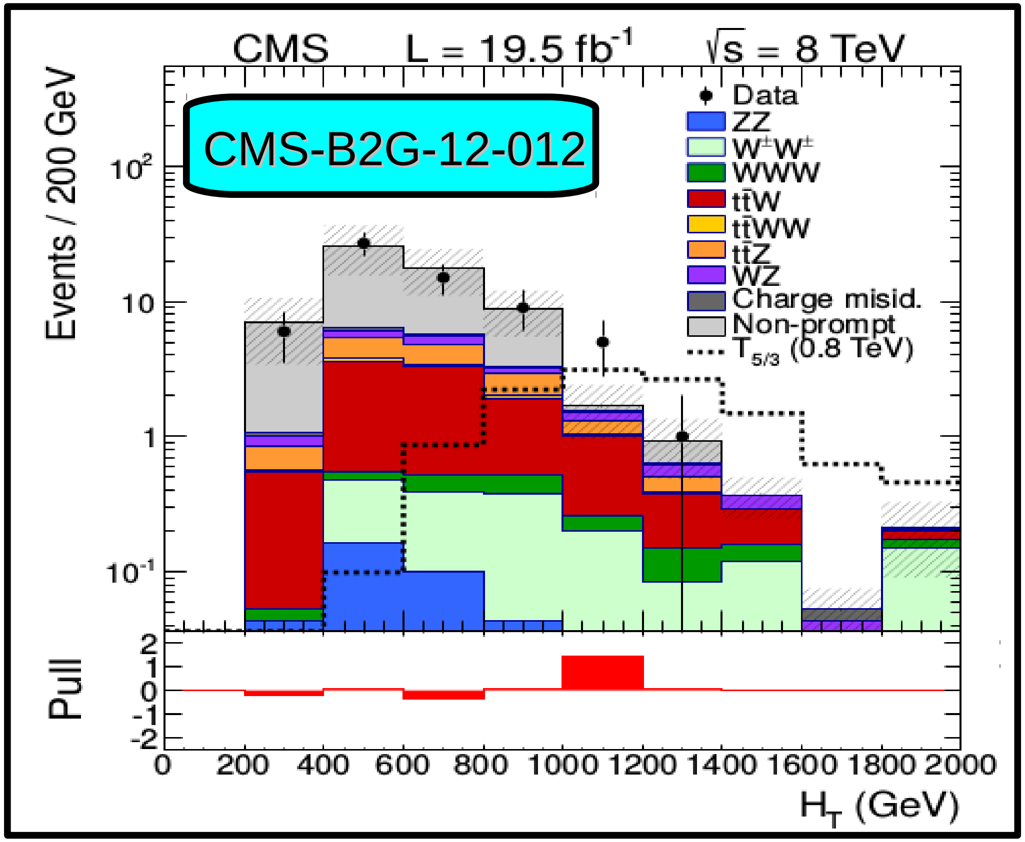
At least two same-sign leptons with $p_T > 30$ GeV, $\Delta R(t,l) > 0.8$
$M_{ll} < 76$ && $M_{ll} > 106$ GeV, $H_T > 900$ GeV
$N_{jet} > 5$ → each AK5 counts 1, W jet counts 2 and Top jet counts 3

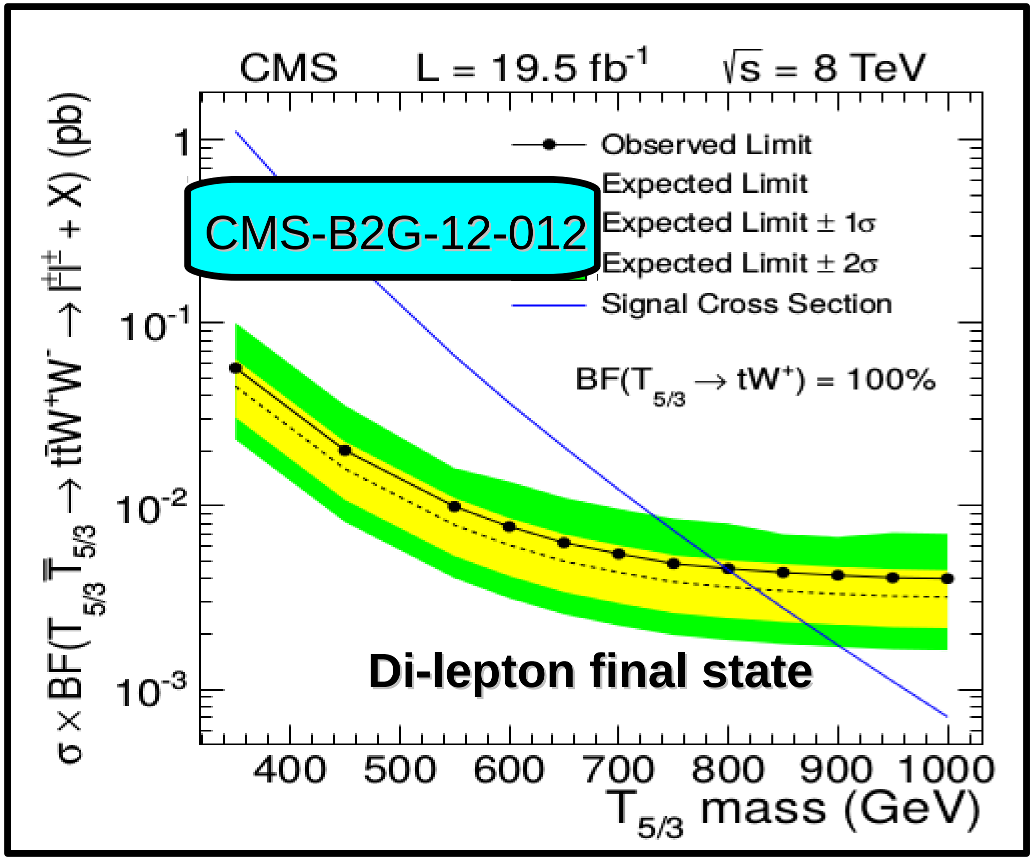
Backgrounds

SM tt → due to large cross section , ttW, ttWW, ttZ, WWW and same-sign WW, WZ, ZZ
Backgrounds due to **mis-charge** reconstruction → from charge **fake rate using data**
Backgrounds due to **fake lepton** reconstruction are estimated from data

T(5/3) quark search: results

- H_T template is used to extract an upper bound on T(5/3) production cross section





Results:

T(5/3) → tt W^+W^- → 2l + X with BR(T(5/3) → tW) = 100% excluded up to 830 GeV

Sensitivity gain at high mass due to sub-structure is ~ [15,20]%

Vector-like quark search

Signal:

Simulated by **MG+Pythia6** with different BR combination

Model:
Charged -1/3
Predicted by little h, composite Higgs, extra dimension

Signature:

b' → tW,bZ or bH, search for b'$\overline{b'}$ decaying in a final state with:
one charged lepton from W, >= 4 jet, >= 1 b-jet, MET

Offline selection:

CA8 used to look for V or H jets, p_T > 200 GeV

Pruned Mass [50,150] GeV, mass drop < 0.4, matched at least with one AK5 jet ΔR < 0.5

4 AK5 jets with p_T > 200,60,40,30 GeV and at least one of them, not matched with CA8, must be b-tagged

Event Categorization: 0, 1 or >= 2 V-jet found

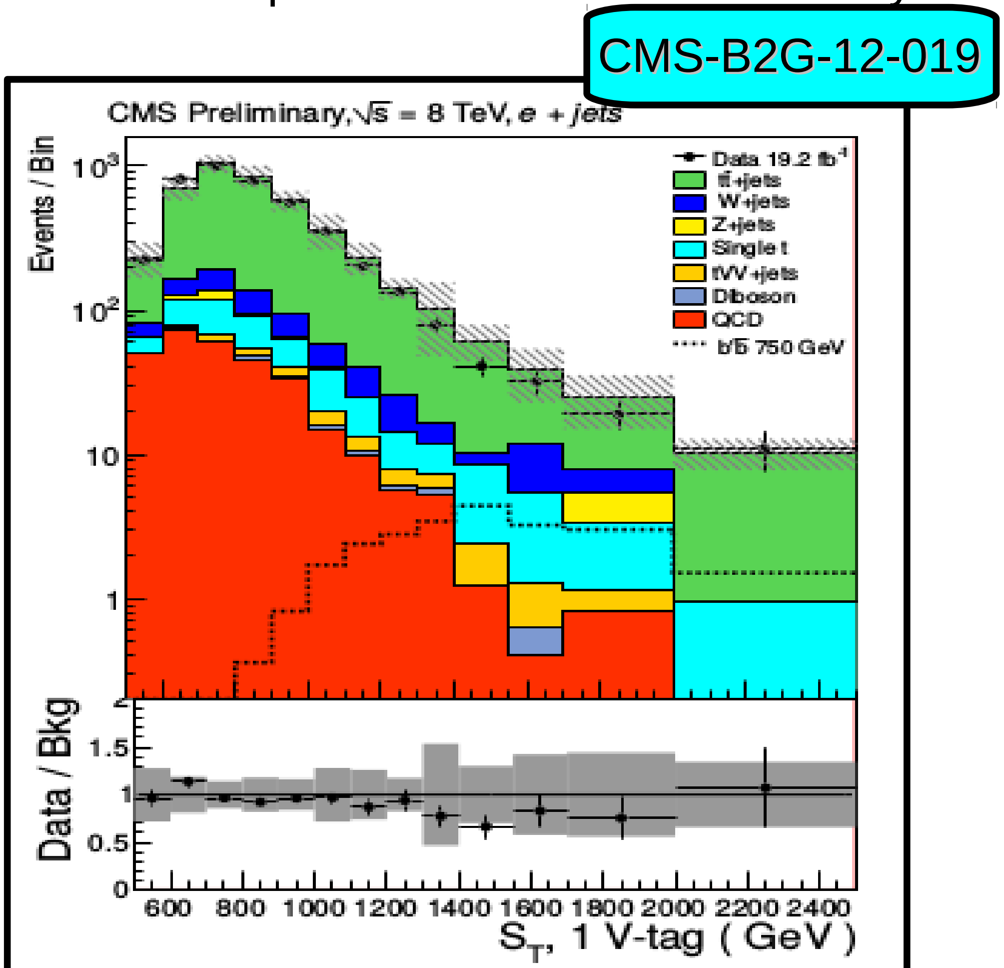
Backgrounds

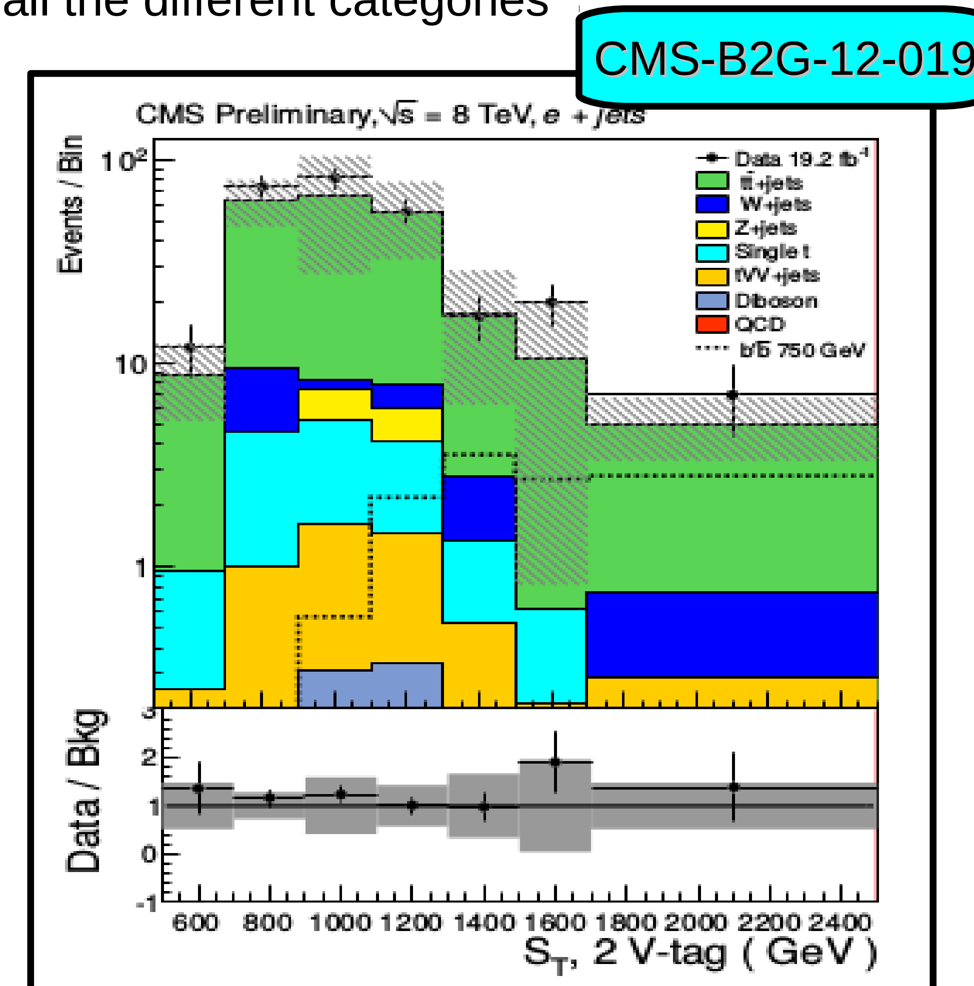
SM $t\bar{t}$, single-top, W+jets, Z+jets, WW/WZ/ZZ → scale factor correction from data applied

QCD multi-jet extracted from **data** fitting the **MET** distribution, obtained inverting **lepton Iso** with a conservative systematics of 100%

Vector-like quark search: spectrum

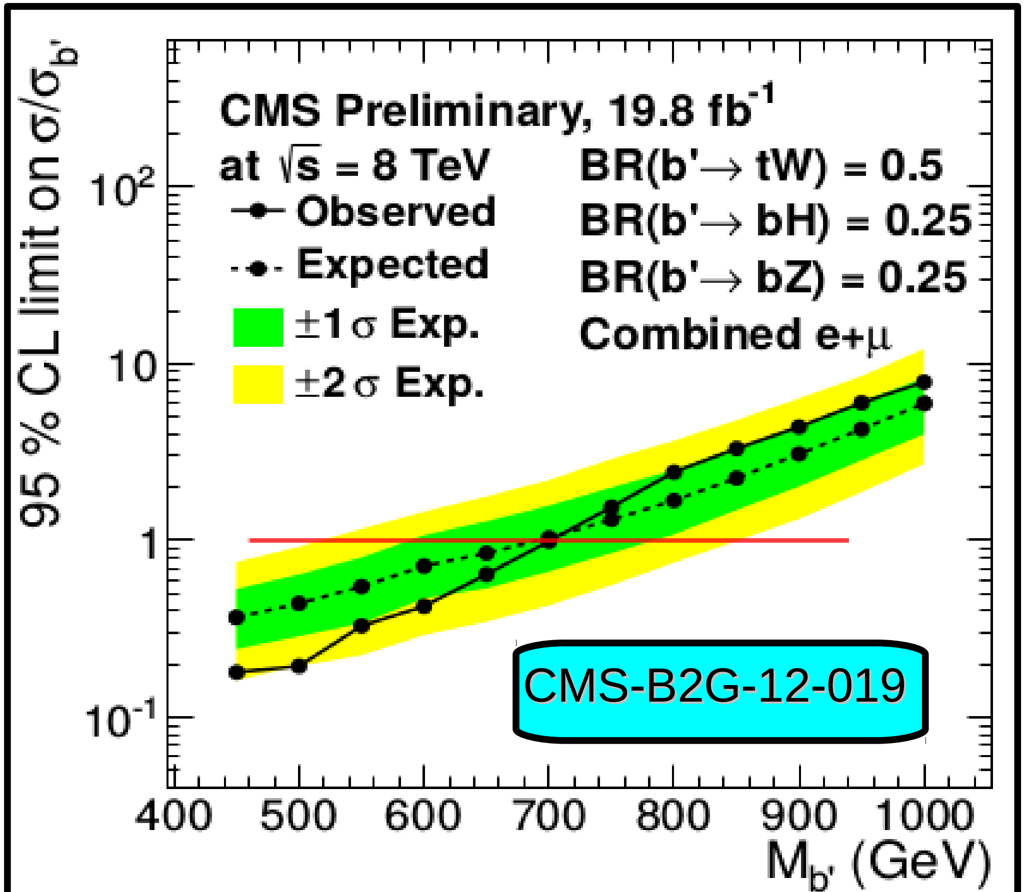
- S_T is the best discriminating variable defined as the scalar sum of AK5 jet, lepton and MET
- Binned template likelihood simultaneously fitting all the different categories

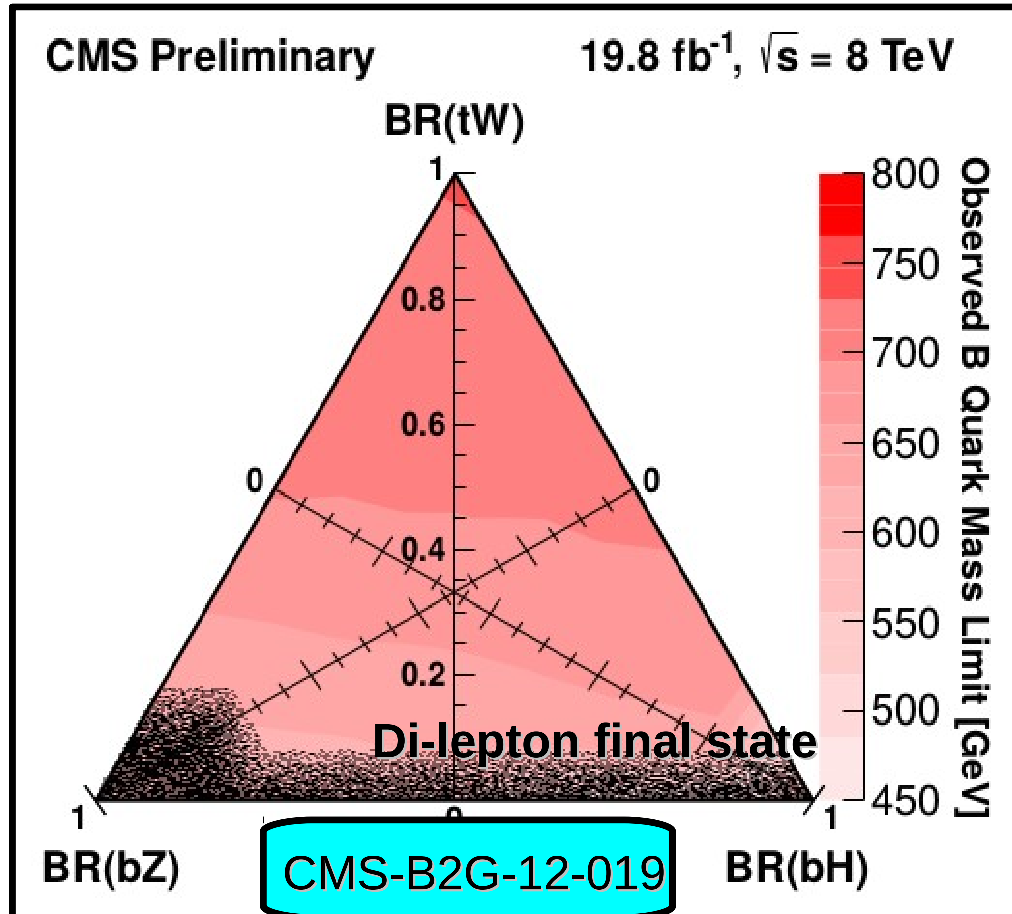




- Main background is SM $t\bar{t}$, all the other are put in only one template in the final fit

Vector-like quark search: results





Results:

Benchmark scenario → observed mass limit at 700 GeV

Analysis more sensitive to tW mode because it goes to 1 and 2 V-tag category

Shaded region means expected sensitivity less than 500 GeV