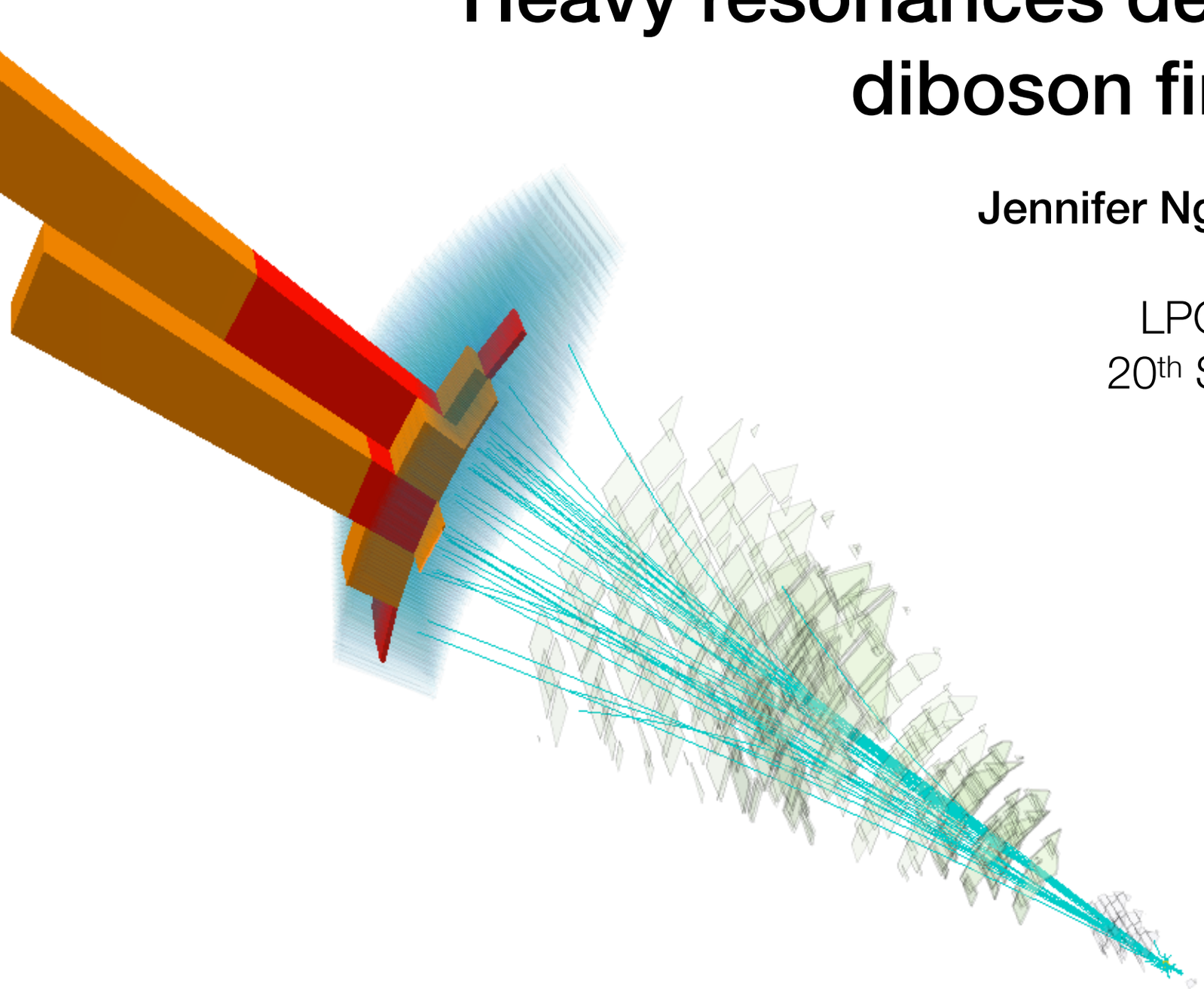
Highlights on searches for new physics with boosted W, Z, and Higgs bosons with CMS

Heavy resonances decaying to diboson final states

Jennifer Ngadiuba (CERN)



LPC Physics Forum
20th September 2018

Why X ➜ diboson?

$M_X > 1$ TeV

The hierarchy problem:

$$\mu^2 = \lambda v^2 = \frac{\lambda}{g^2} 4M_W^2 \sim 10^4\ \mathrm{GeV}^2 \ll M_{Pl} \sim 10^{38}\ \mathrm{GeV}^2$$

Possible solutions (among others):

Composite Higgs: the Higgs is not elementary → new spin-1 resonances (W', Z')
Large/warped extra-dimensions: gravity is strong at electroweak scale → new spin-0/2 resonances (graviton/radion)
2HDM: introduce a second scalar doublet → new spin-0/1 resonances

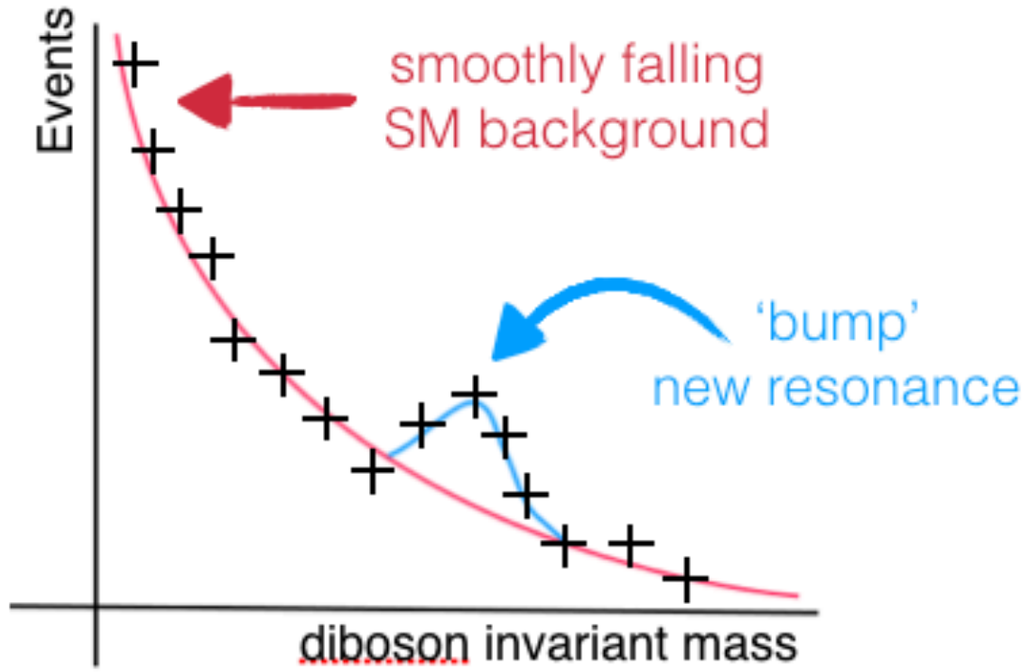
How do we observe
X ➜ diboson at LHC?

$M_X > 1$ TeV

X produced through **gluon-fusion, qq annihilation or VBF**

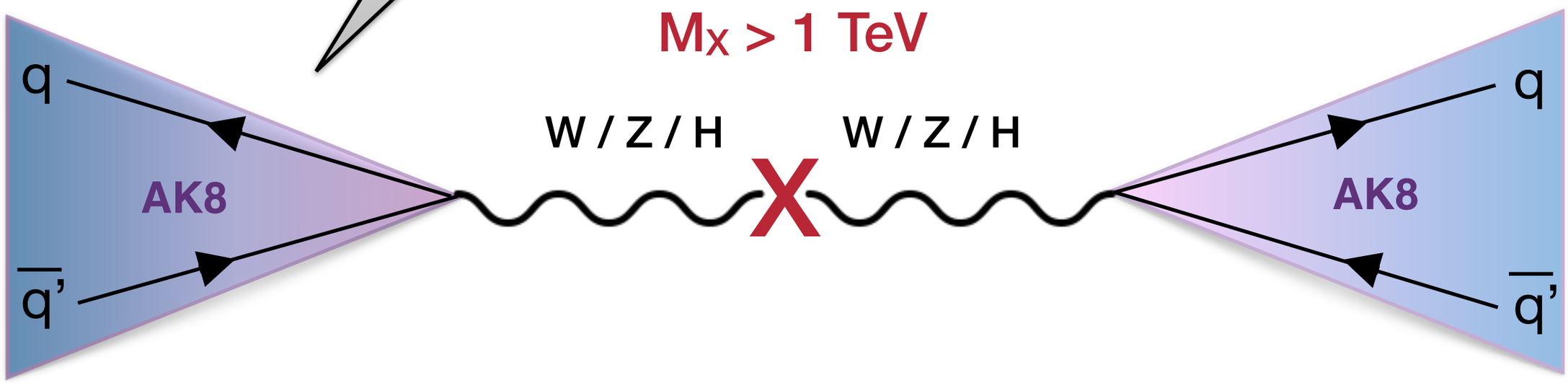
search for narrow resonances,
natural width < experimental resolution

reconstruct the two bosons and build invariant mass ⇒ **search for "bumps"**



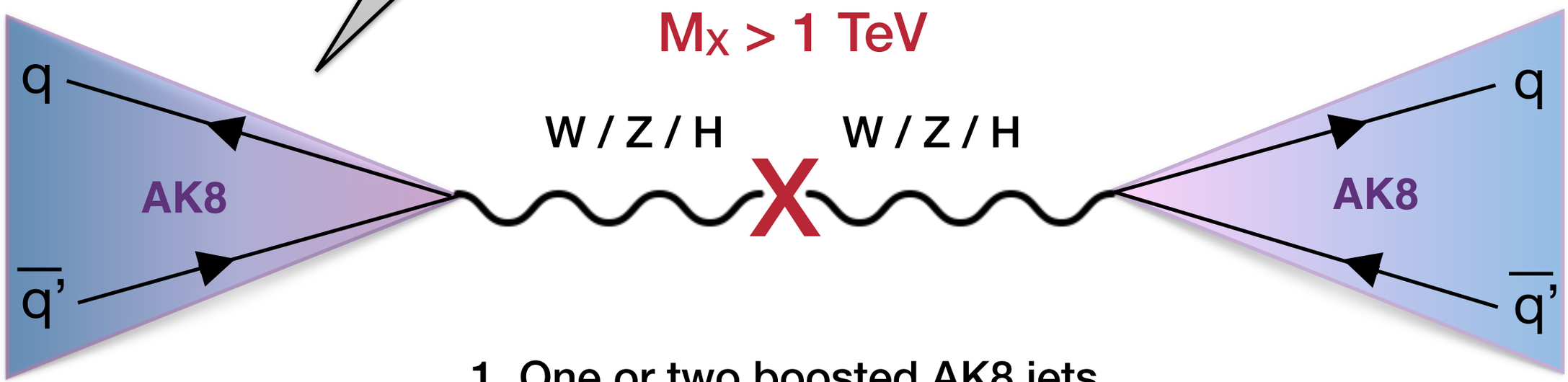
Boson reconstruction?

Search in all bosons decay modes



Boson reconstruction?

Search in all bosons decay modes



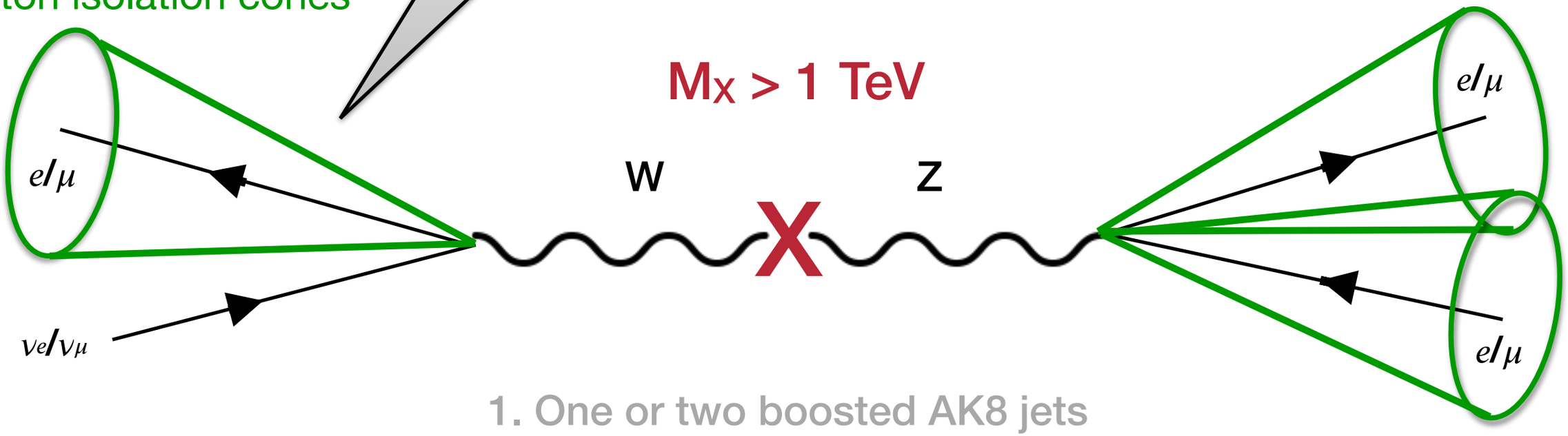
1. One or two boosted AK8 jets

boson tagging exploiting jet substructure

Boson reconstruction?

Search in all bosons decay modes

Lepton isolation cones



1. One or two boosted AK8 jets

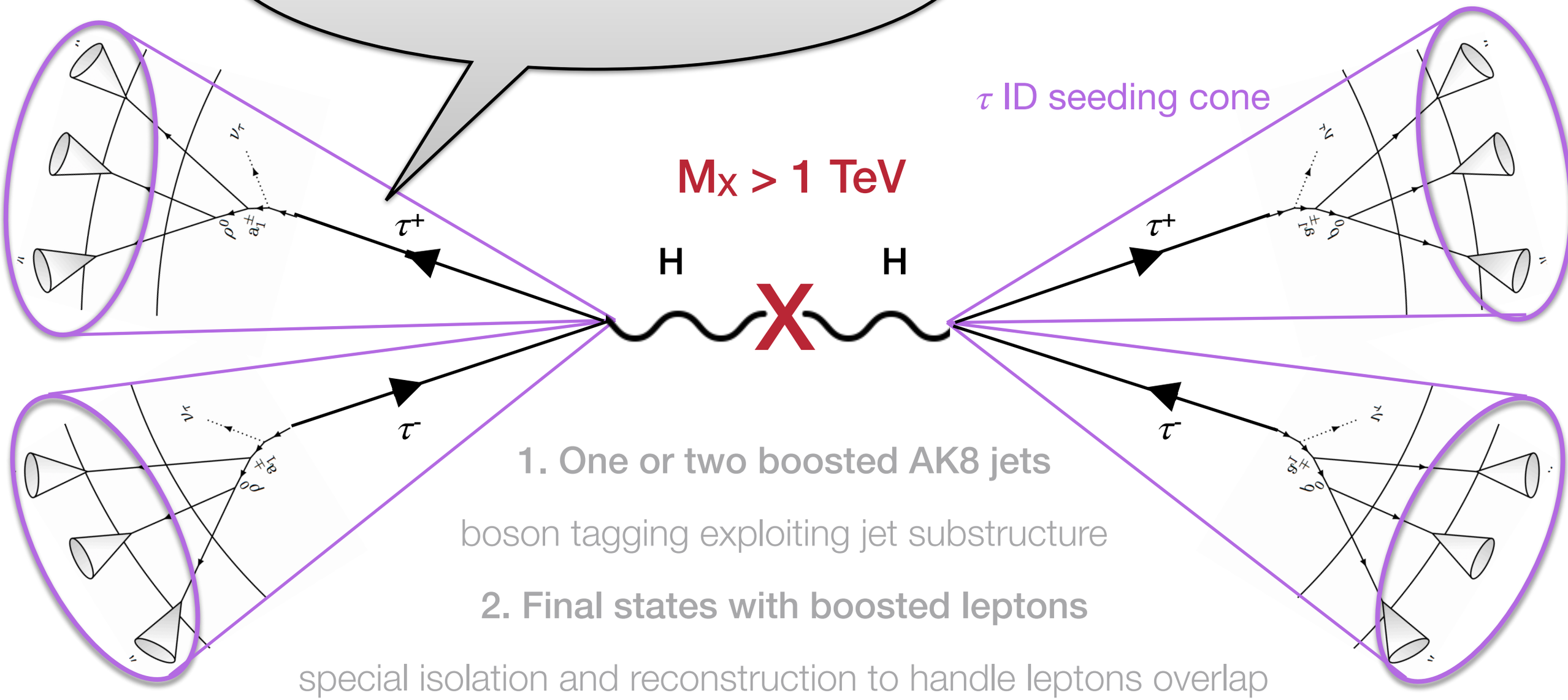
boson tagging exploiting jet substructure

2. Final states with boosted leptons

special isolation and reconstruction to handle leptons overlap

Boson reconstruction?

Search in all bosons decay modes



1. One or two boosted AK8 jets

boson tagging exploiting jet substructure

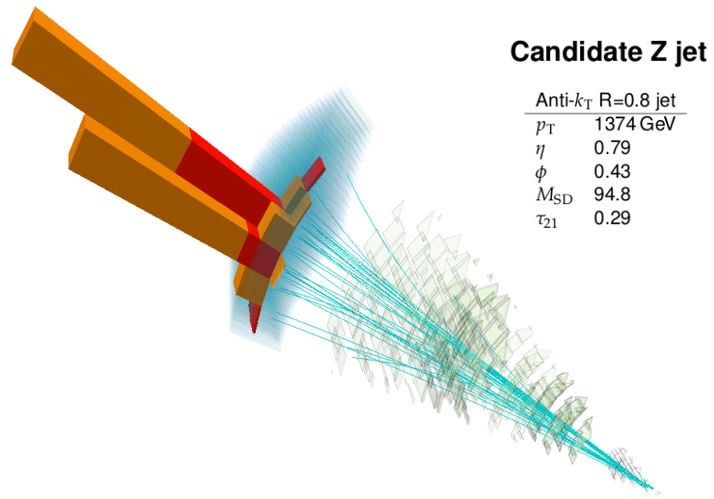
2. Final states with boosted leptons

special isolation and reconstruction to handle leptons overlap

3. Final states with boosted taus

dedicated tau reconstruction and identification

Hadronic boson identification



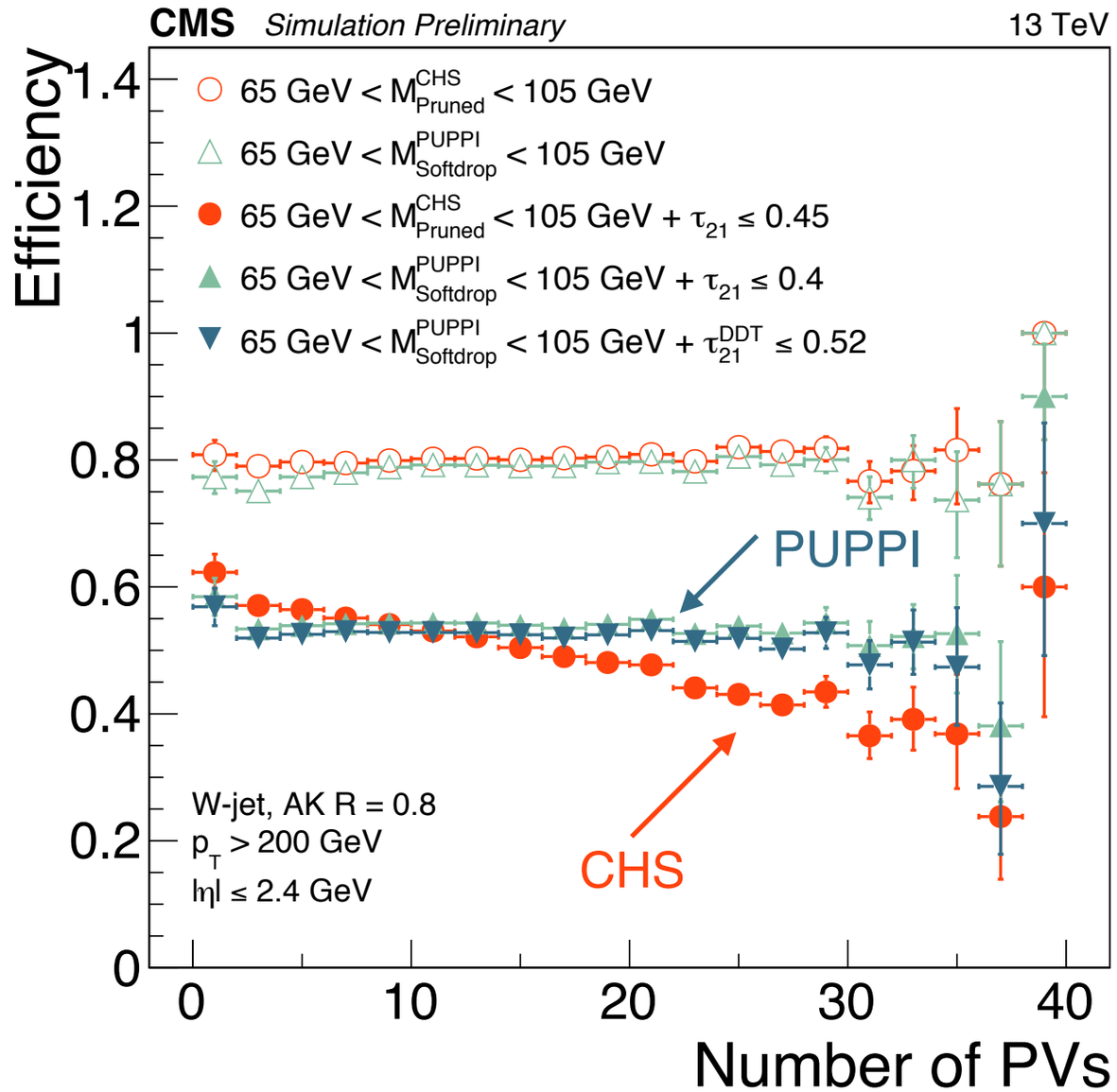
1. Get rid of pile-up

(JHEP 1410 (2014) 59)

PileUp Per Particle Identification, PUPPI
(CMS "default" for AK8 jets)

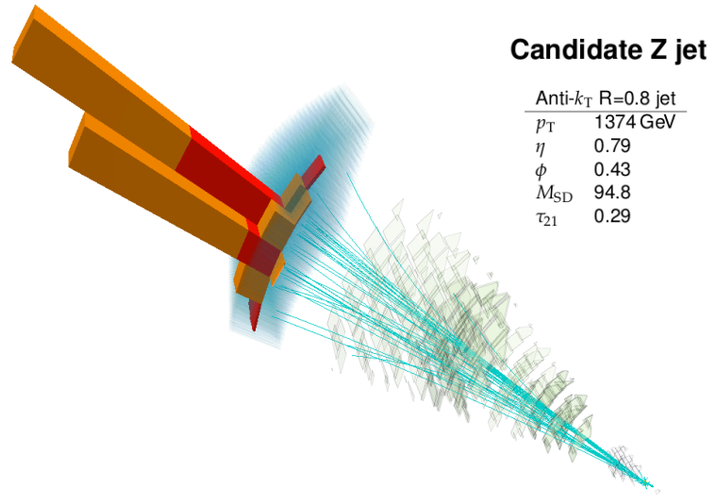
Charged Hadron Subtraction (CHS):
Removes charged particles associated to PU vertex, but no neutral particles
Correct jet quantities for neutral PU

PUPPI:
Calculate weight for each particle based on the shape of charged PU
Remove particles rather than correct jet quantities



CMS-PAS-JME-16-003

Hadronic boson identification



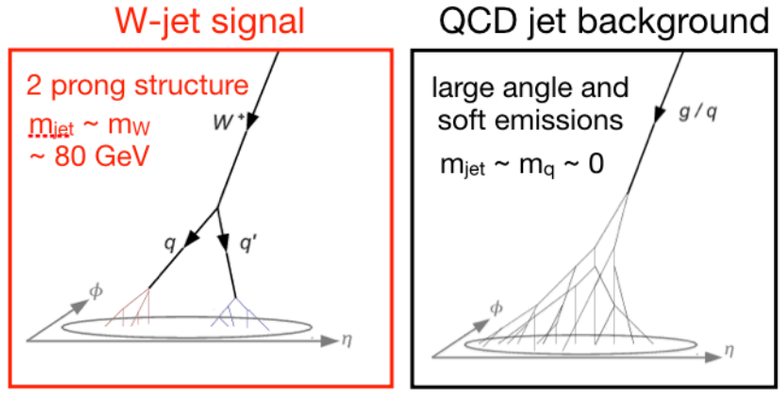
1. Get rid of pile-up (JHEP 1410 (2014) 59)

PileUp Per Particle Identification, PUPPI
(CMS "default" for AK8 jets)

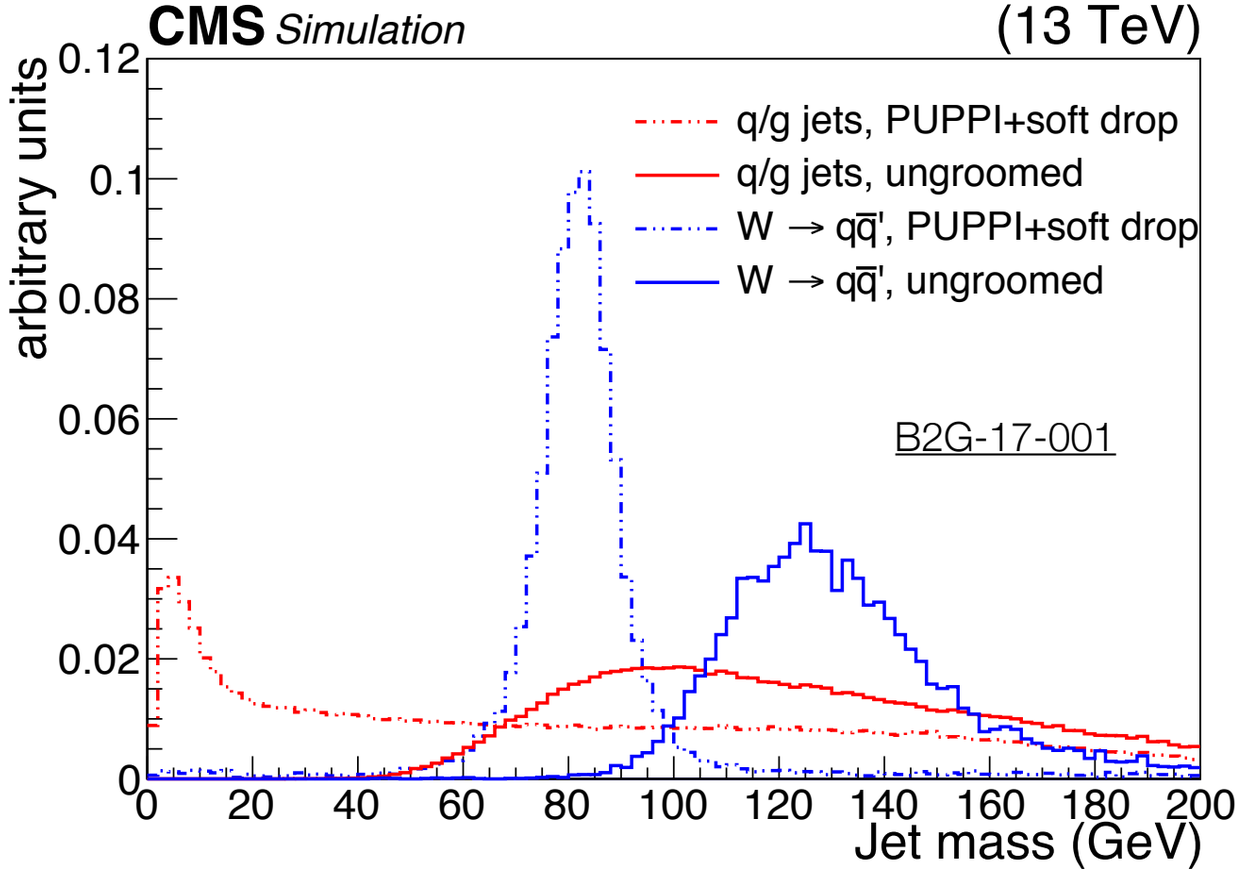
JHEP 1405 (2014) 146

2. Reconstruct mass with jet grooming

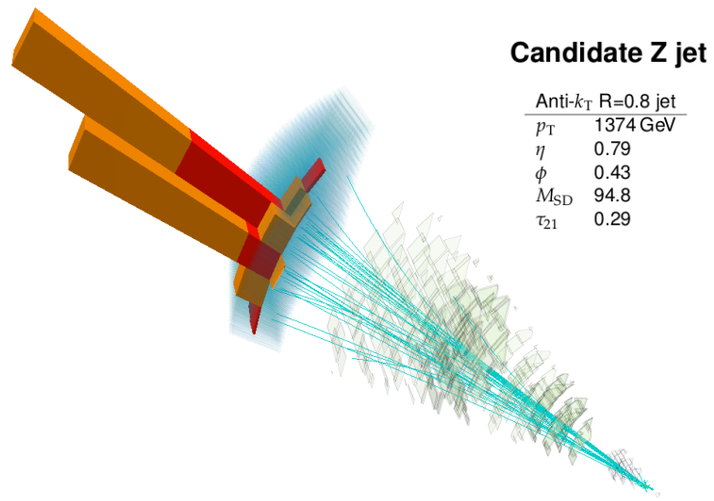
modified Mass Drop Tagger
(softdrop β=0, z=0.1)



Softdrop:
filters out large angle and soft emissions
→ push the mass of the QCD jet to zero!



Hadronic boson identification



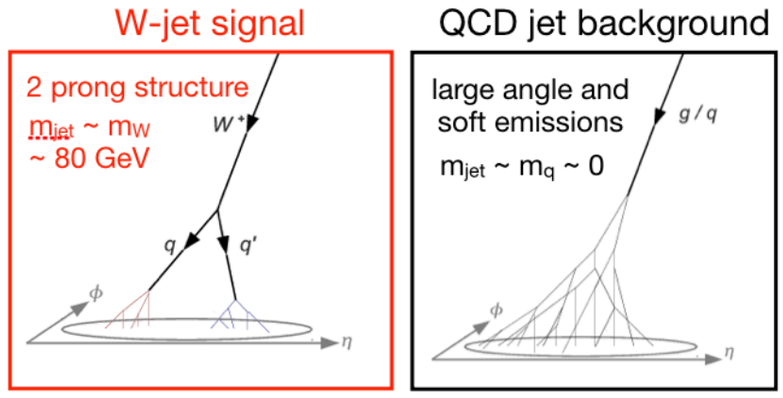
1. Get rid of pile-up (JHEP 1410 (2014) 59)

PileUp Per Particle Identification, PUPPI
(CMS "default" for AK8 jets)

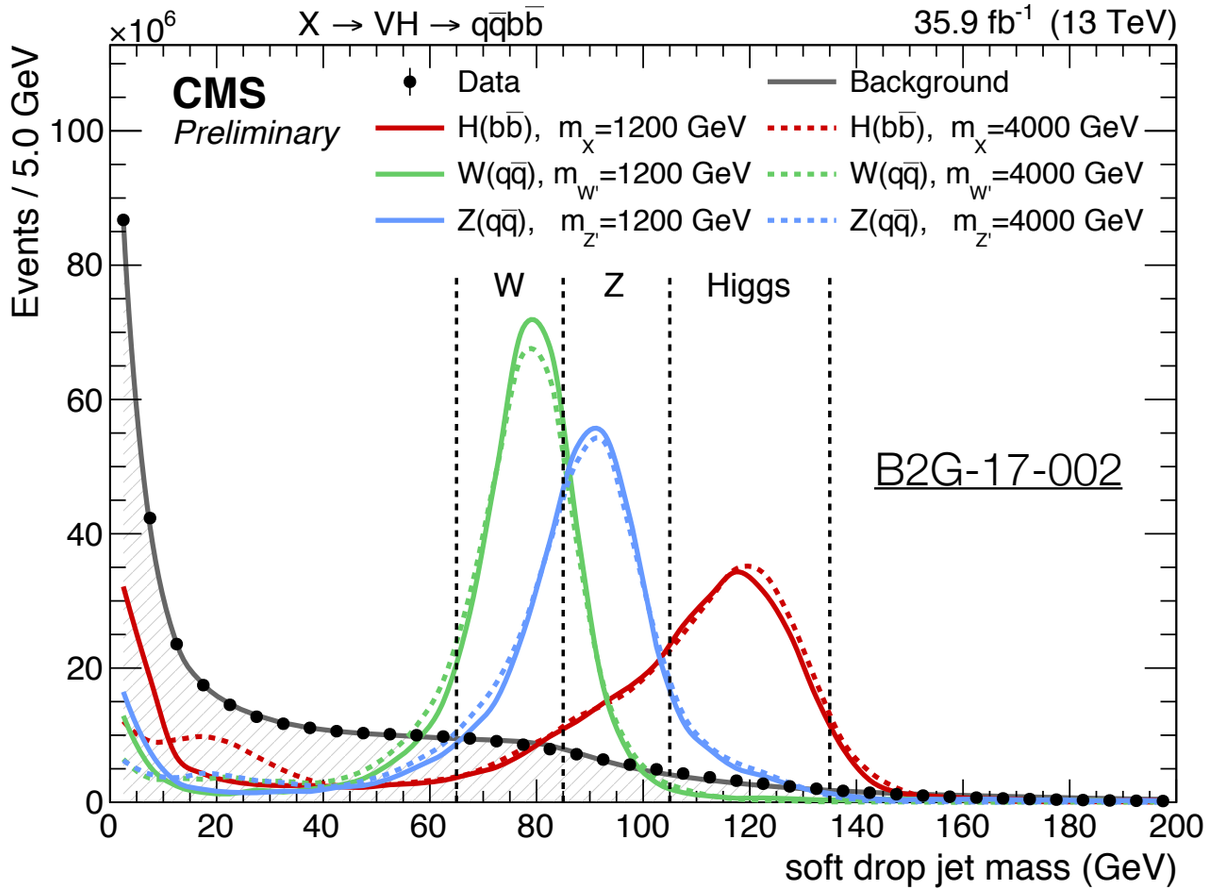
JHEP 1405 (2014) 146

2. Reconstruct mass with jet grooming

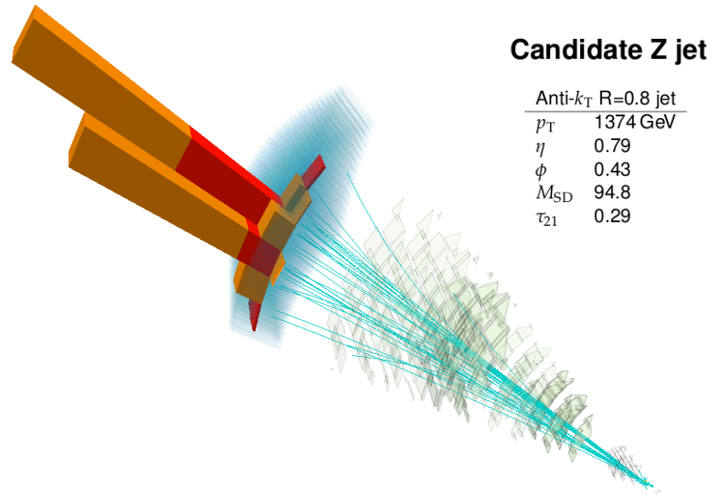
modified Mass Drop Tagger
(softdrop β=0, z=0.1)



! Orthogonal mass windows
Important for combining results, mass resolution is key for choice of groomer



Hadronic boson identification



1. Get rid of pile-up (JHEP 1410 (2014) 59)

PileUp Per Particle Identification, PUPPI
(CMS "default" for AK8 jets)

JHEP 1405 (2014) 146

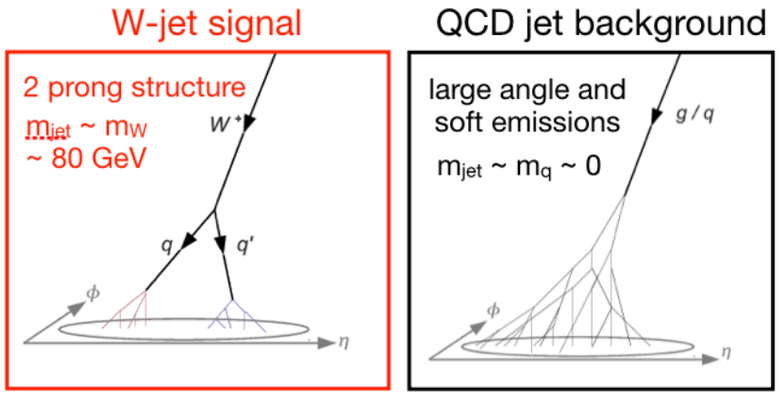
2. Reconstruct mass with jet grooming

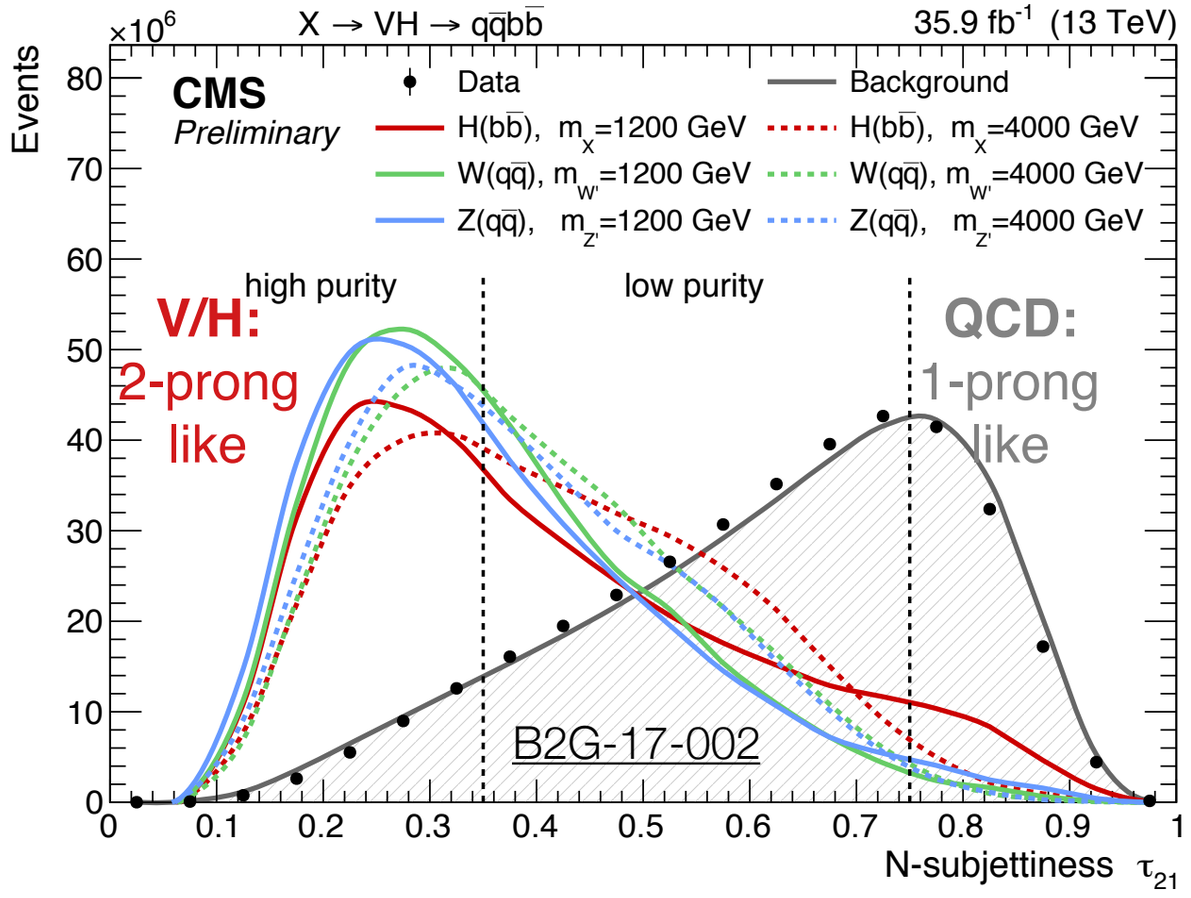
modified Mass Drop Tagger
(softdrop β=0, z=0.1)

3. Tag by resolving jet substructure

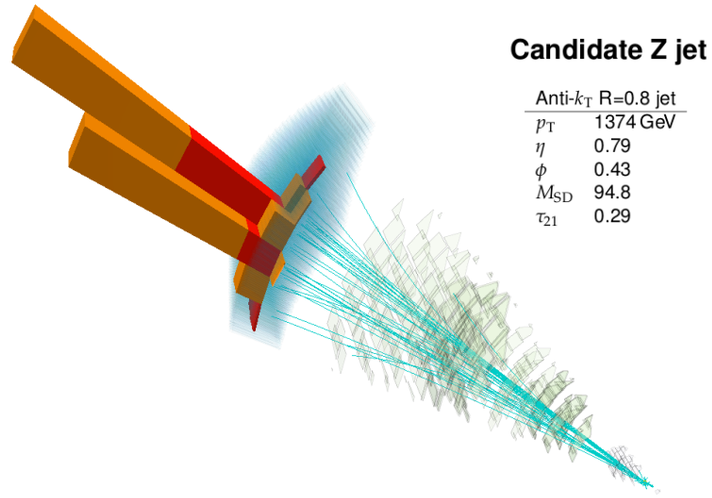
n-subjetiness ratio τ_{21}

N-sujettiness:
Decluster the jet in N subjets and evaluate compatibility with N subjets hypothesis





Hadronic boson identification



1. Get rid of pile-up (JHEP 1410 (2014) 59)

PileUp Per Particle Identification, PUPPI
(CMS "default" for AK8 jets)

JHEP 1405 (2014) 146

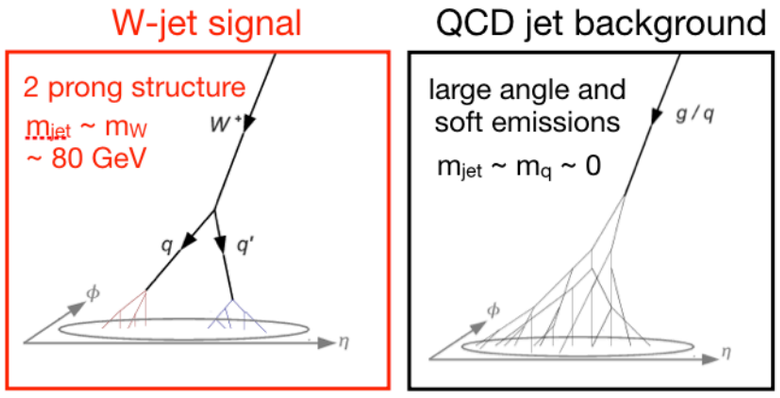
2. Reconstruct mass with jet grooming

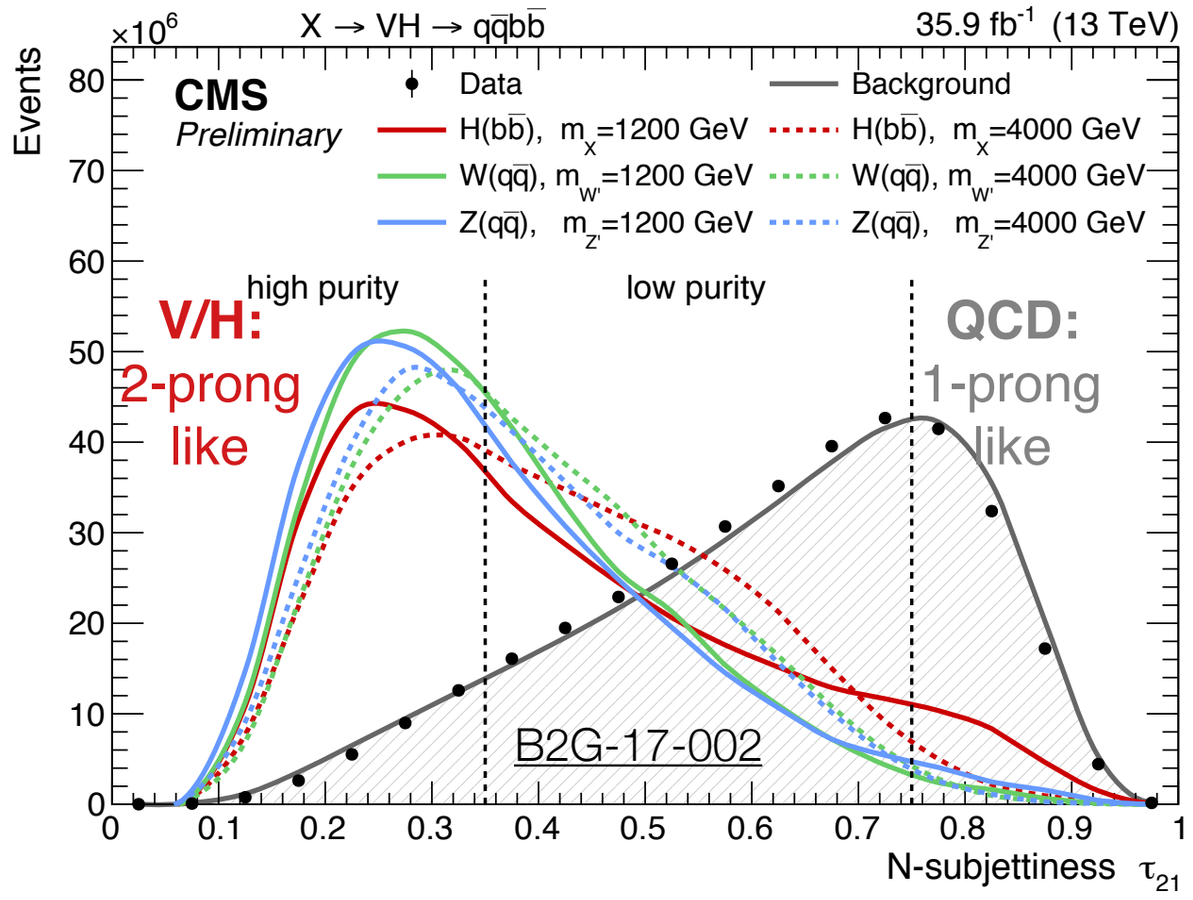
modified Mass Drop Tagger
(softdrop β=0, z=0.1)

3. Tag by resolving jet substructure

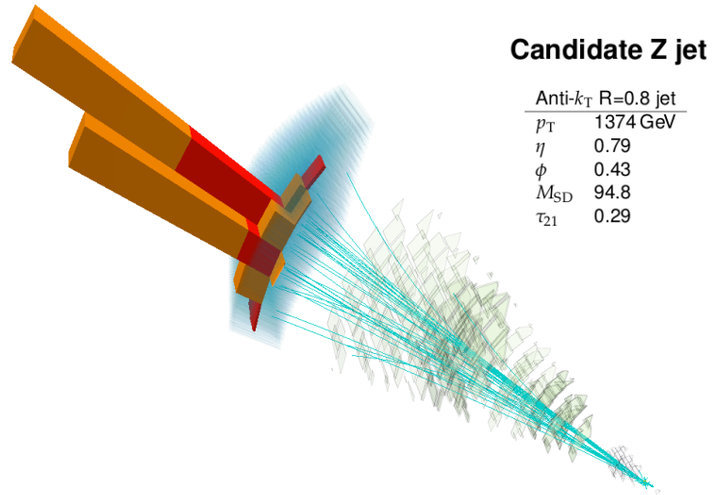
n-subjetiness ratio τ_{21}

! Two τ_{21} categories:
high-purity: best S/B
low-purity: high efficiency





Hadronic boson identification



1. Get rid of pile-up (JHEP 1410 (2014) 59)

PileUp Per Particle Identification, PUPPI
(CMS "default" for AK8 jets)

JHEP 1405 (2014) 146

2. Reconstruct mass with jet grooming

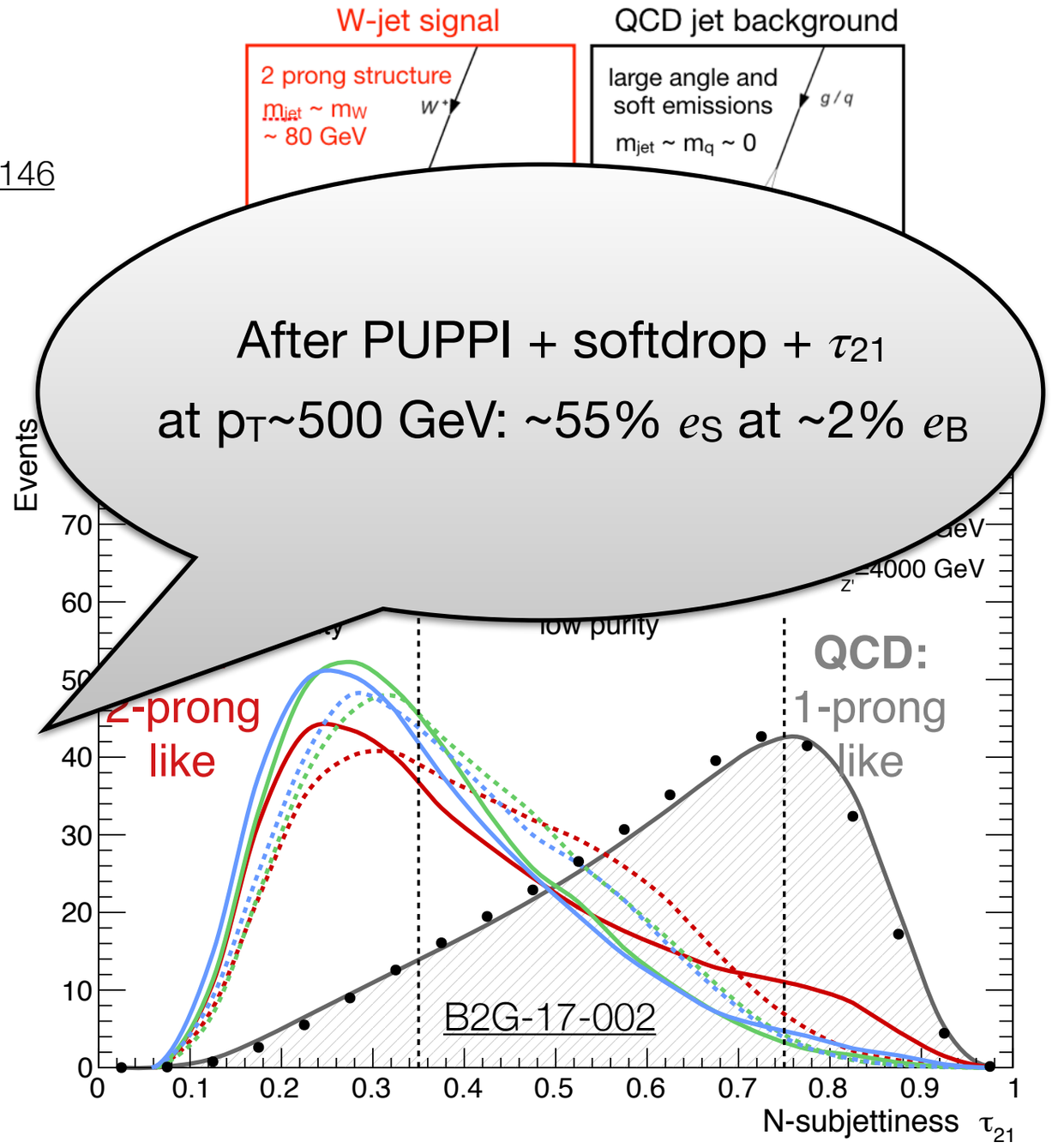
modified Mass Drop Tagger
(softdrop β=0, z=0.1)

3. Tag by resolving jet substructure

n-subjetiness ratio τ_{21}

! Two τ_{21} categories:

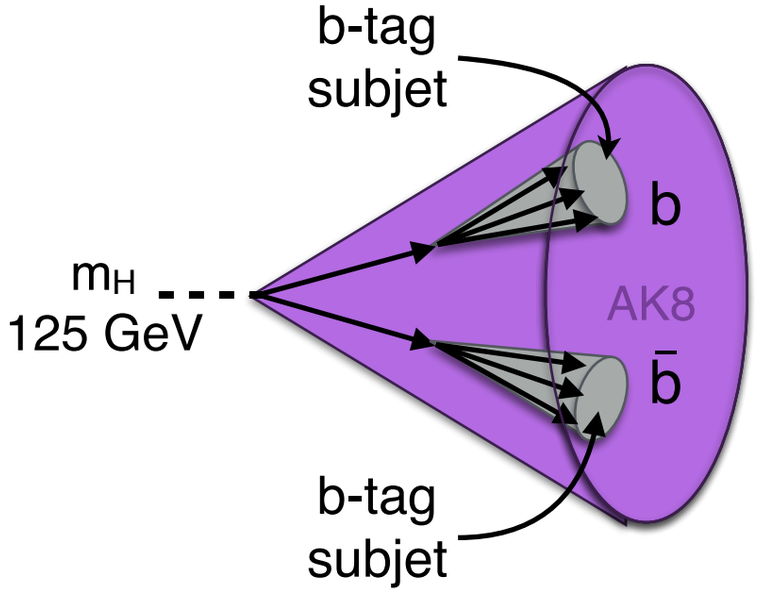
high-purity: best S/B
low-purity: high efficiency



H→bb identification

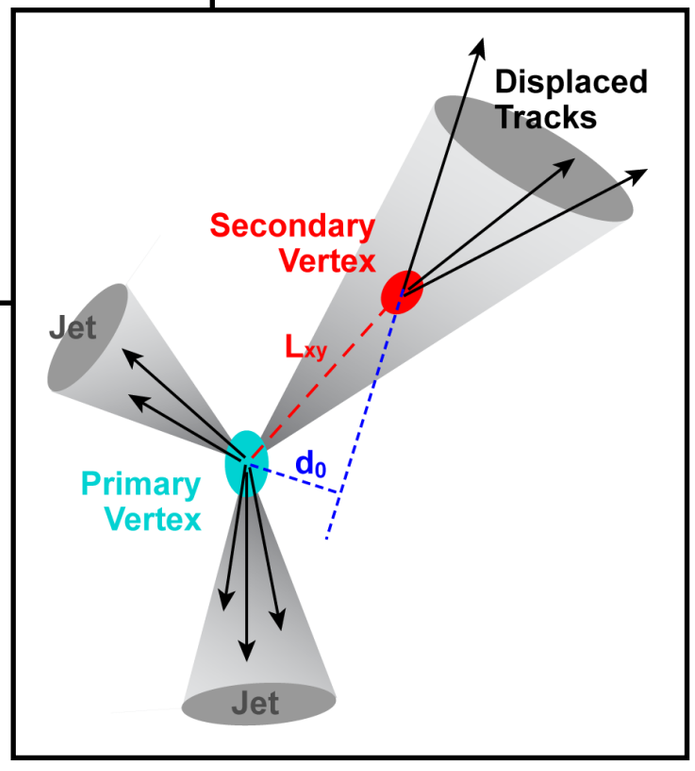
On top of mass and substructure, take advantage of b-tagging. Three methods:

1. tag softdrop subjets (R=0.2) with CSV(*)



(*) default CMS b-tagger (Combined Secondary Vertex), combines b-hadron specific information into ML algo (BDT or DNN):

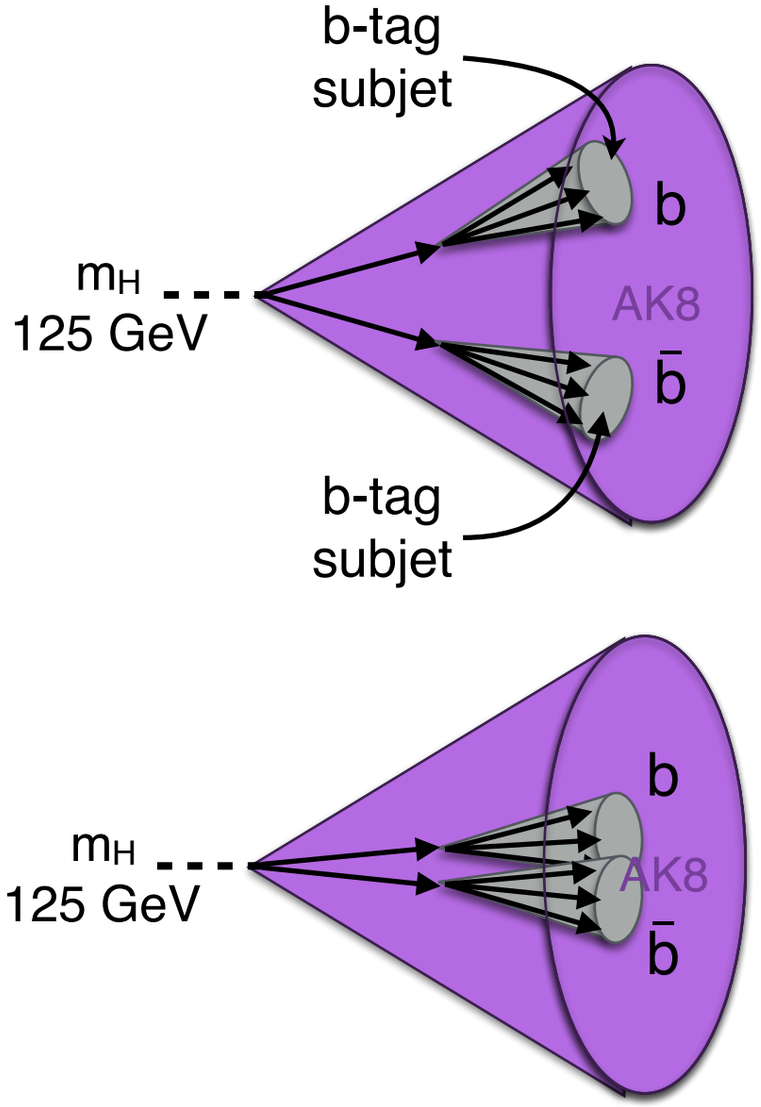
Large lifetime of b-hadrons
Displaced vertices/tracks
Large impact parameters
Non-isolated leptons



H→bb identification

On top of mass and substructure, take advantage of b-tagging. Three methods:

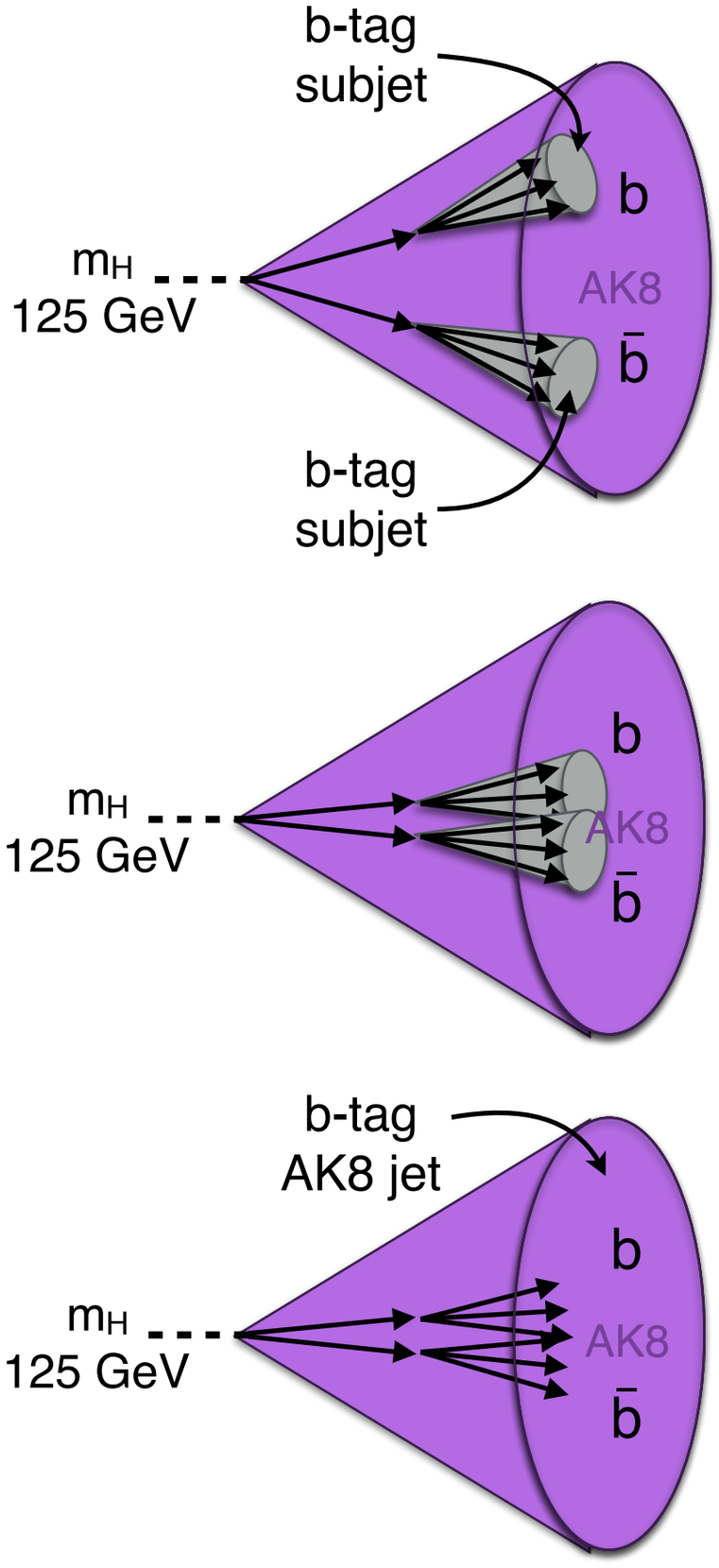
1. tag softdrop subjets (R=0.2) with CSV
2. if boost is high, subjets overlap. Track-sharing leads to performance drop



H→bb identification

On top of mass and substructure, take advantage of b-tagging. Three methods:

1. tag softdrop subjets (R=0.2) with CSV
2. if boost is high, subjets overlap. Track-sharing leads to performance drop
   → b-tag AK8 jet

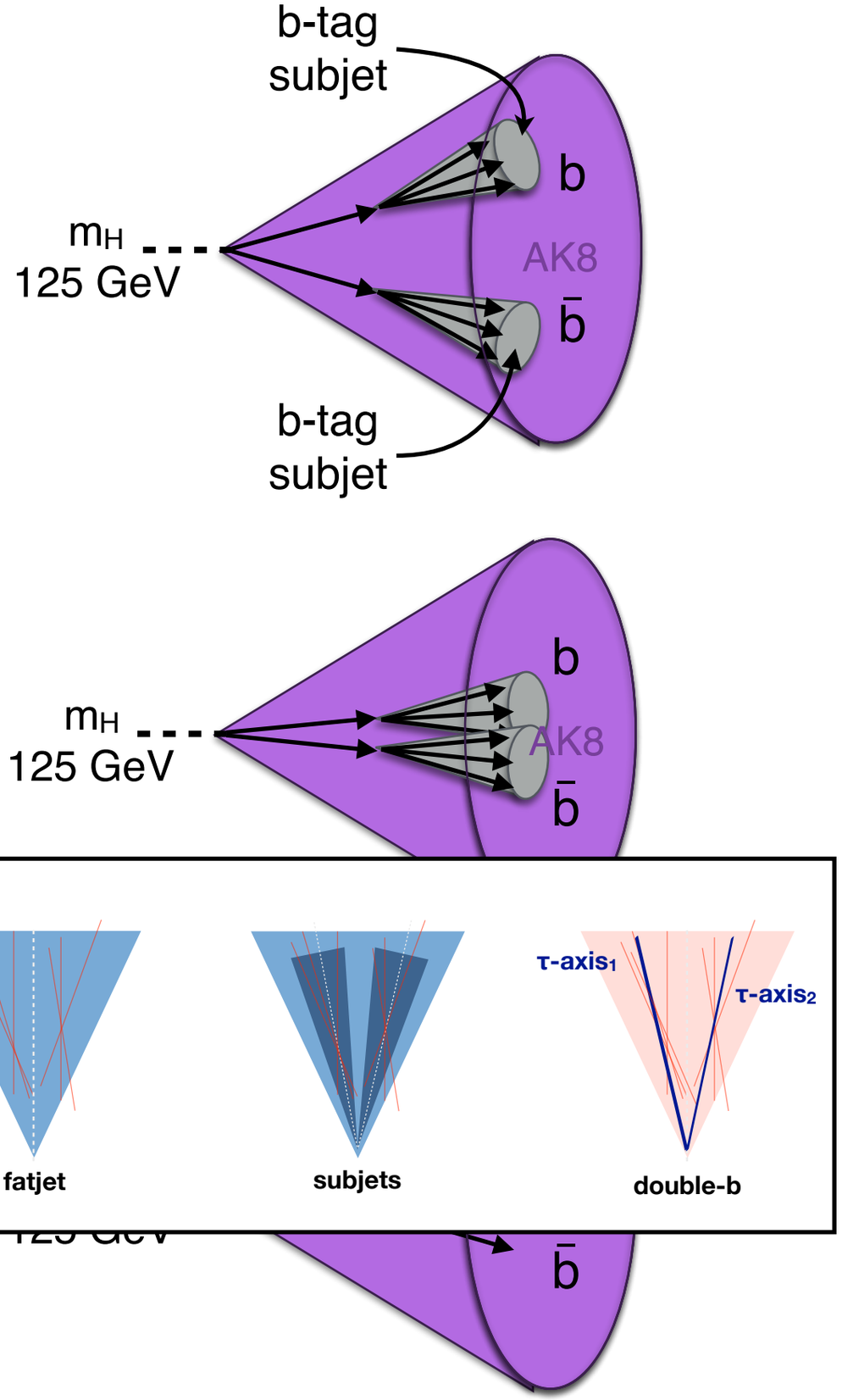


H→bb identification

On top of mass and substructure, take advantage of b-tagging. Three methods:

1. tag softdrop subjets (R=0.2) with CSV
2. if boost is high, subjets overlap. Track-sharing leads to performance drop
3. Alternative: double-b dedicated tagger based on MVA

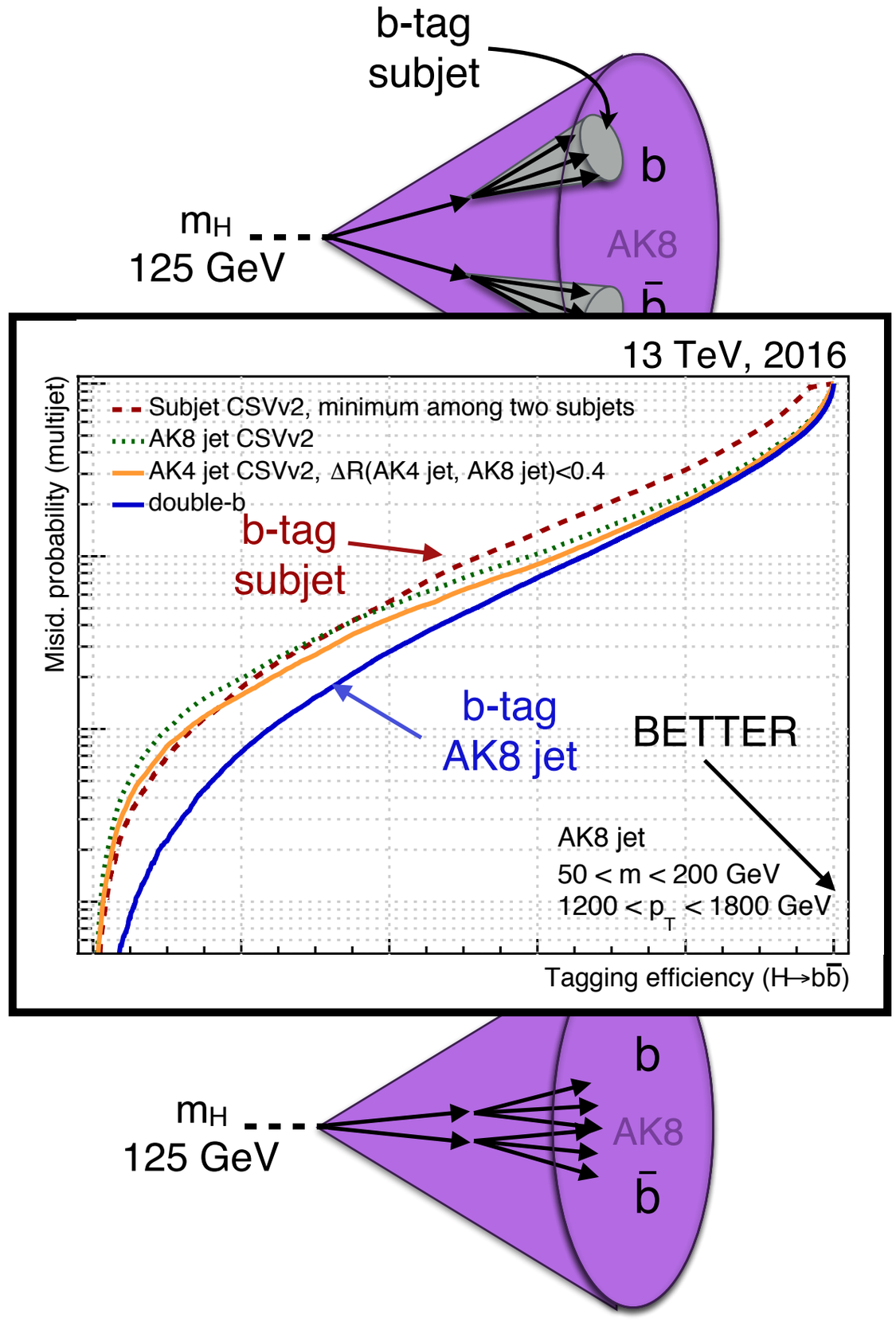
combines b-hadron specific information with substructure observables
exploit correlation between b-hadron flight directions and the two N-subjettiness axis directions



H→bb identification

On top of mass and substructure, take advantage of b-tagging. Three methods:

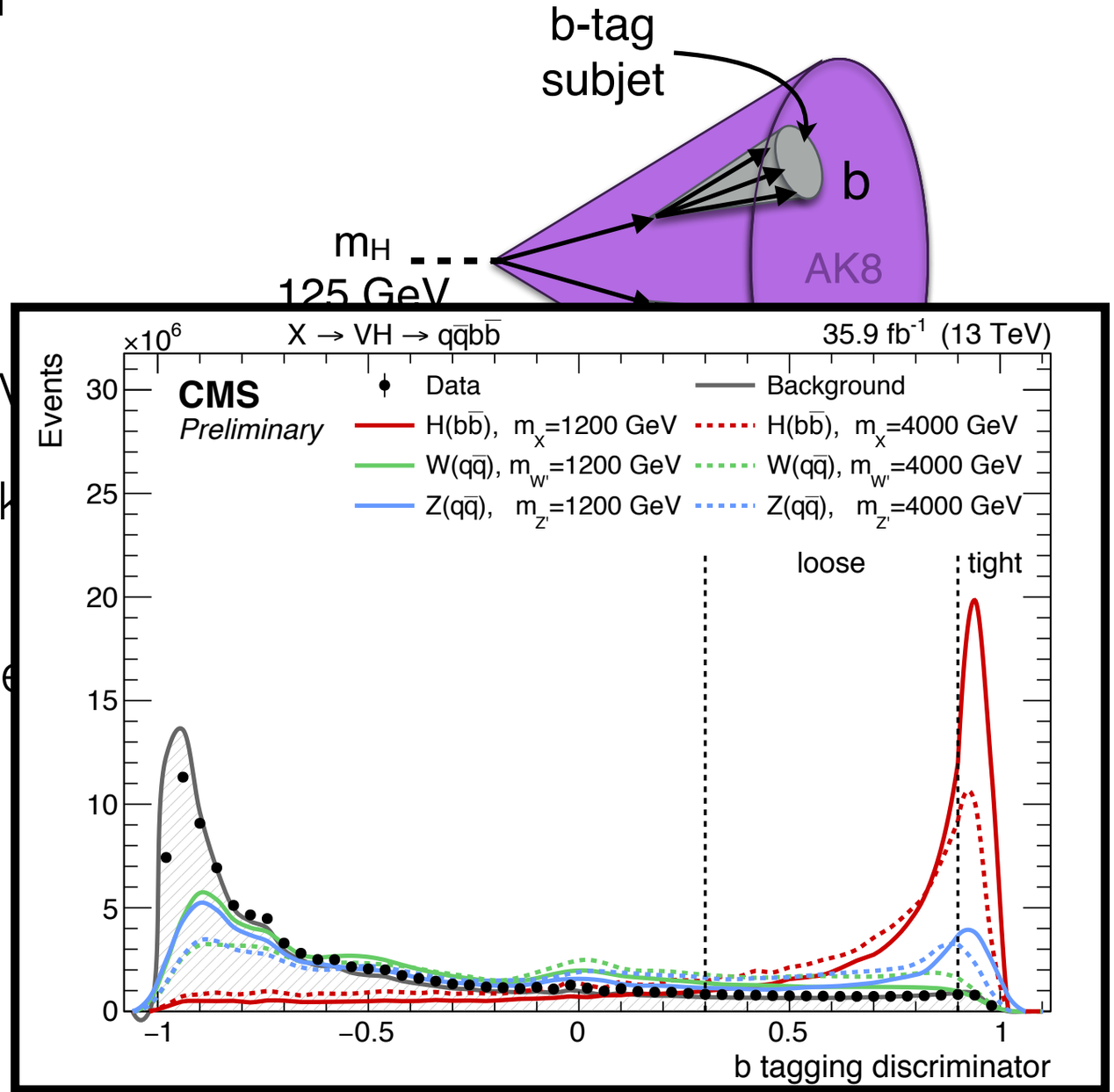
1. tag softdrop subjets (R=0.2) with CSV
2. if boost is high, subjets overlap. Track-sharing leads to performance drop
3. Alternative: double-b dedicated tagger based on MVA



H→bb identification

On top of mass and substructure, take advantage of b-tagging. Three methods:

1. tag softdrop subjets (R=0.2) with CSV
2. if boost is high, subjets overlap. Track sharing leads to performance drop
3. Alternative: double-b dedicated tagger based on MVA



Discriminator (MVA output) working points for jet p_T = 1 TeV

	Loose	Tight
Tagging efficiency	80 %	35 %
Mistag rate	8 %	0,6 %

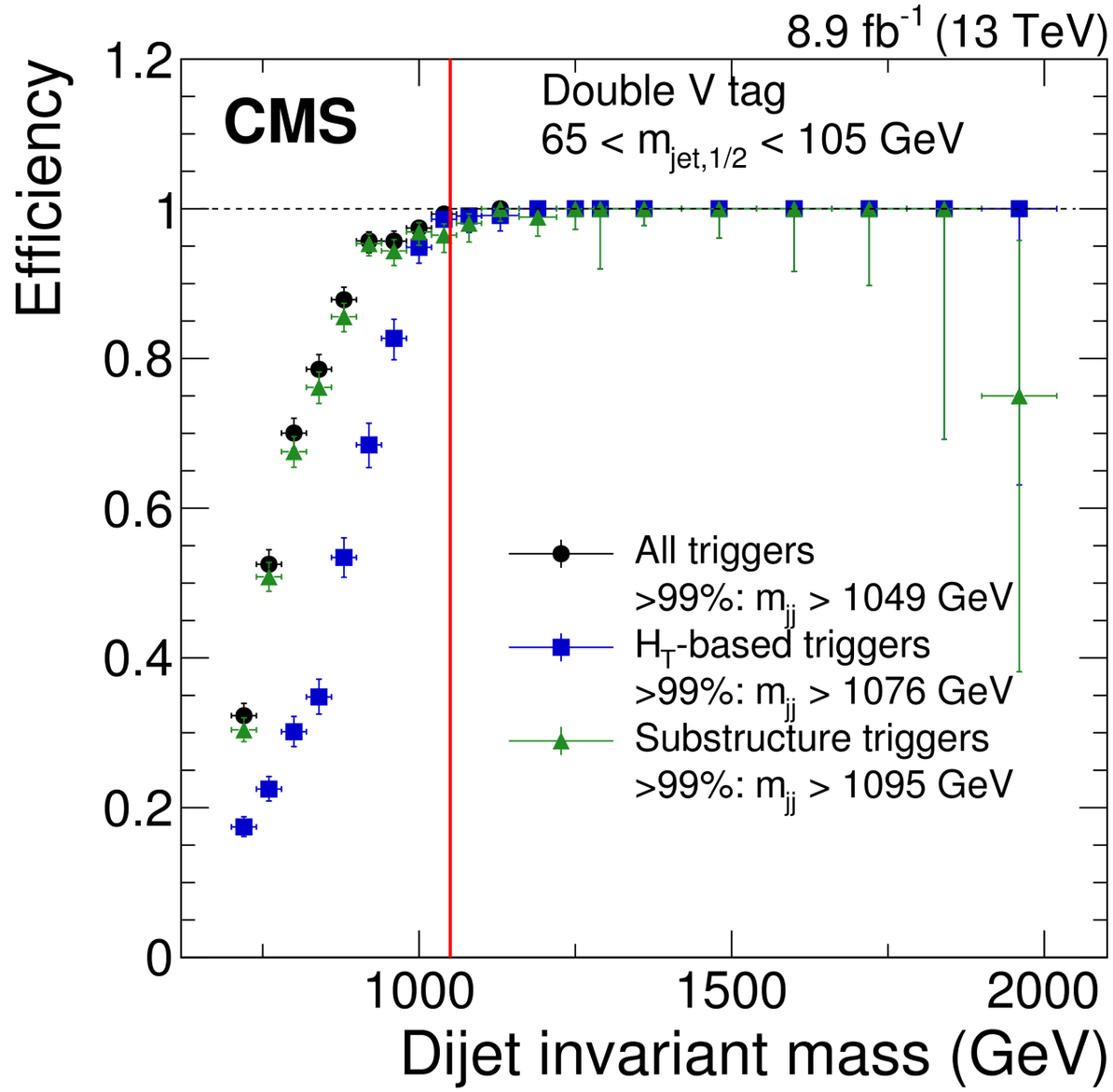
Triggering on boosted jets

Grooming algorithms at HLT lowers m_{jj} trigger thresholds

cut on jet trimmed mass (slightly less aggressive than mMDT used offline) of 30/50 GeV

fully efficient at offline softdrop mass of ~50 GeV

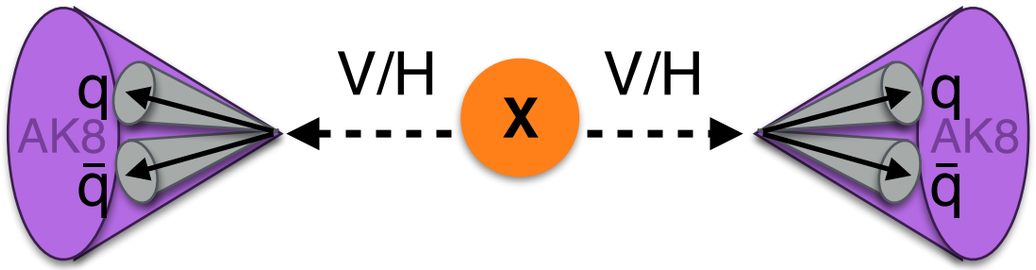
As of 2018, new double-b tag + trimmed mass trigger further lowering thresholds!



All-hadronic final states

BR(H→bb) ~ 57%
BR(V→qq) ~ 69%

Dijet final states: X→VV, VH or HH (V = W, Z)

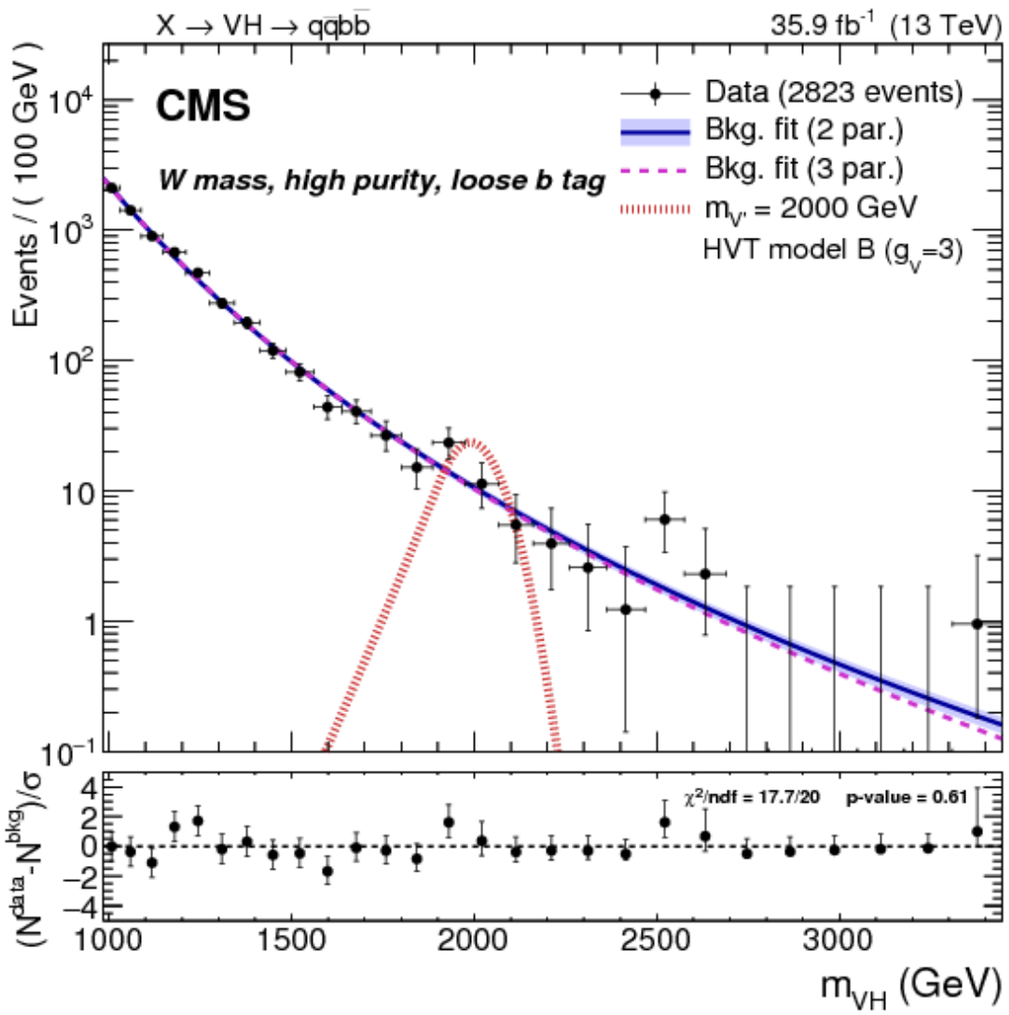


The pros:

largest branching fraction (good at high M_X where background is low)

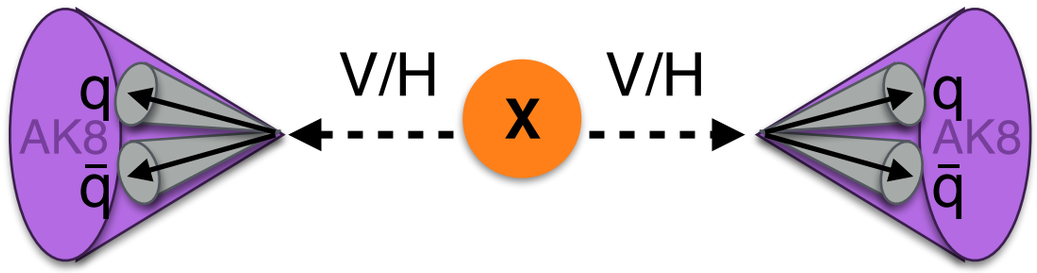
simple and robust: background (QCD) estimated from fit to data with smoothly falling function

B2G-17-002 [EPJC 77 (2017) 636]



All-hadronic final states

BR(H→bb) ~ 57%
BR(V→qq) ~ 69%



Dijet final states: X→VV, VH or HH (V = W, Z)

The pros:

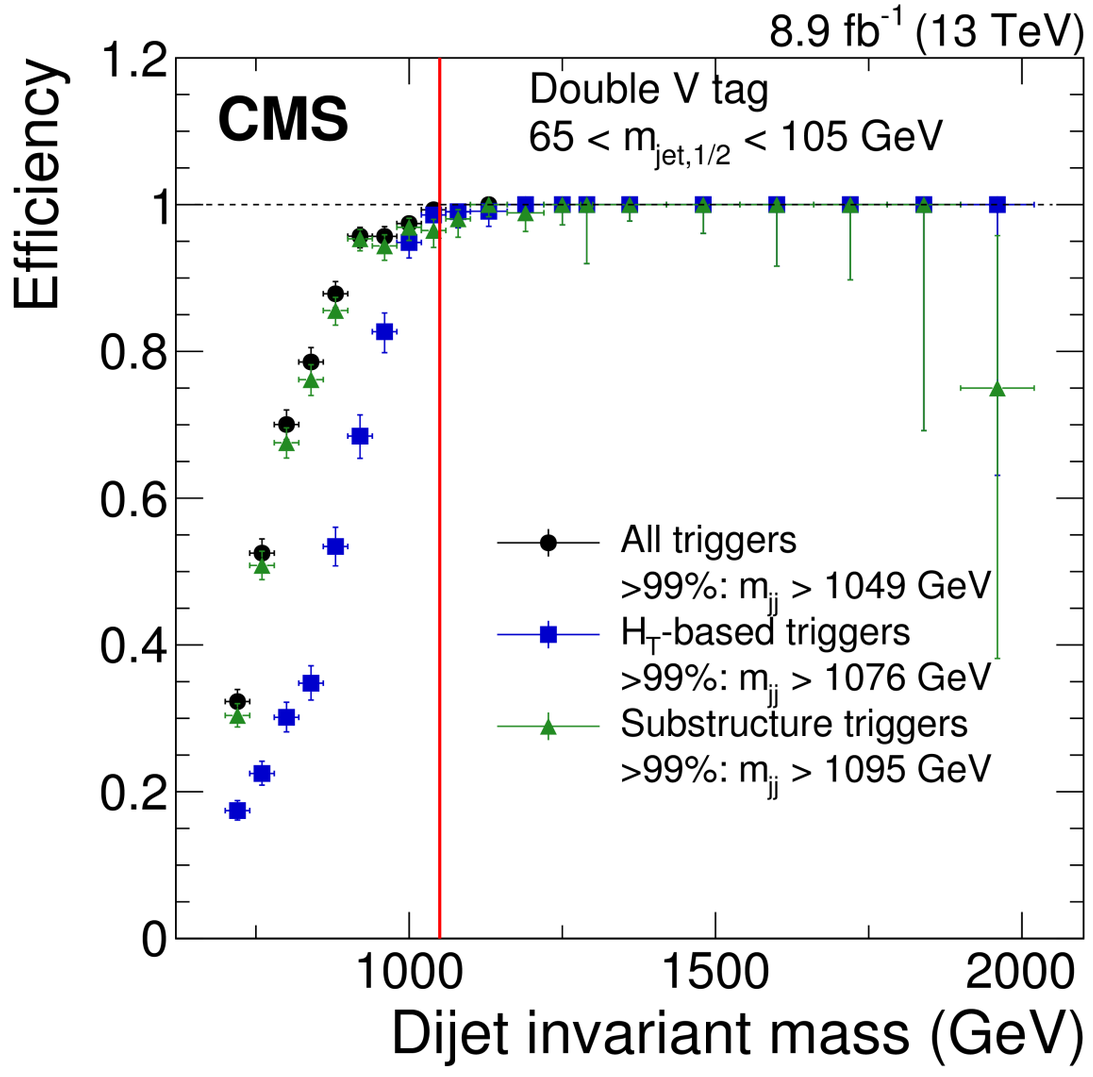
largest branching fraction (good at high M_X where background is low)

simple and robust: background (QCD) estimated from fit to data with smoothly falling function

The cons:

trigger limited (need smoothly falling m_{jj})

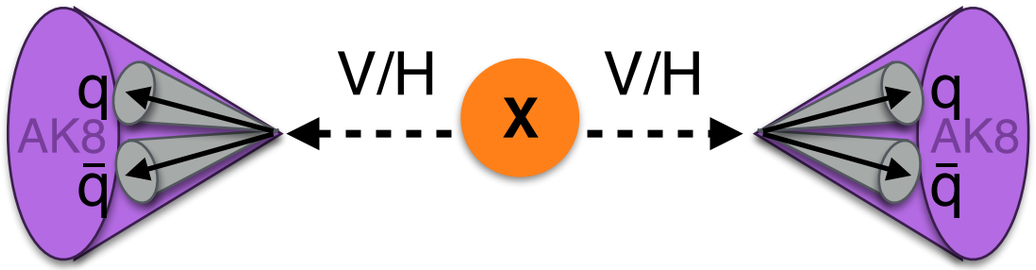
B2G-17-001 [Phys. Rev. D 97, 072006 (2018)]



All-hadronic final states

BR(H→bb) ~ 57%
BR(V→qq) ~ 69%

Dijet final states: X→VV, VH or HH (V = W, Z)



The pros:

largest branching fraction (good at high M_X where background is low)

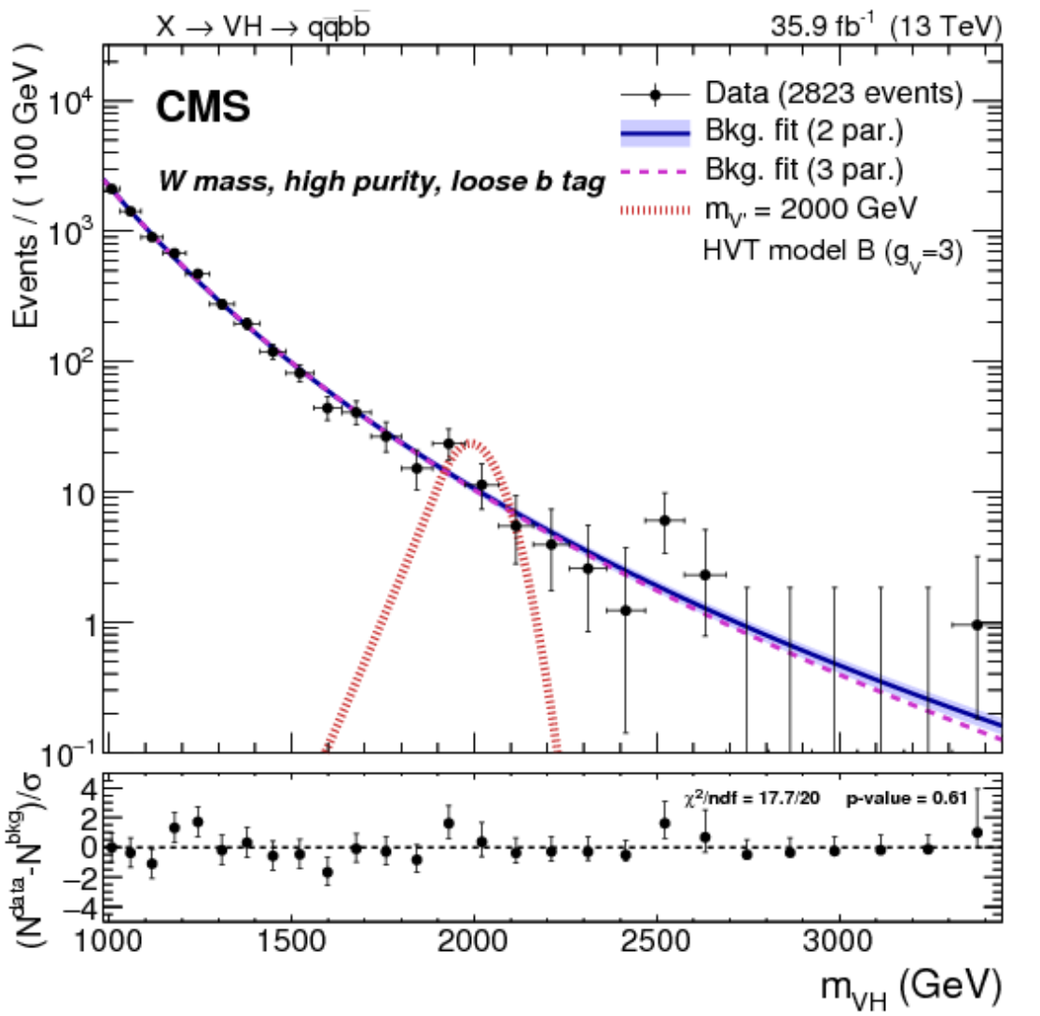
simple and robust: background (QCD) estimated from fit to data with smoothly falling function

The cons:

trigger limited (need smoothly falling m_{jj})

background-only fit sensitive to upward fluctuation in the tails → extensive bias test typically conducted

B2G-17-001 [Phys. Rev. D 97, 072006 (2018)]



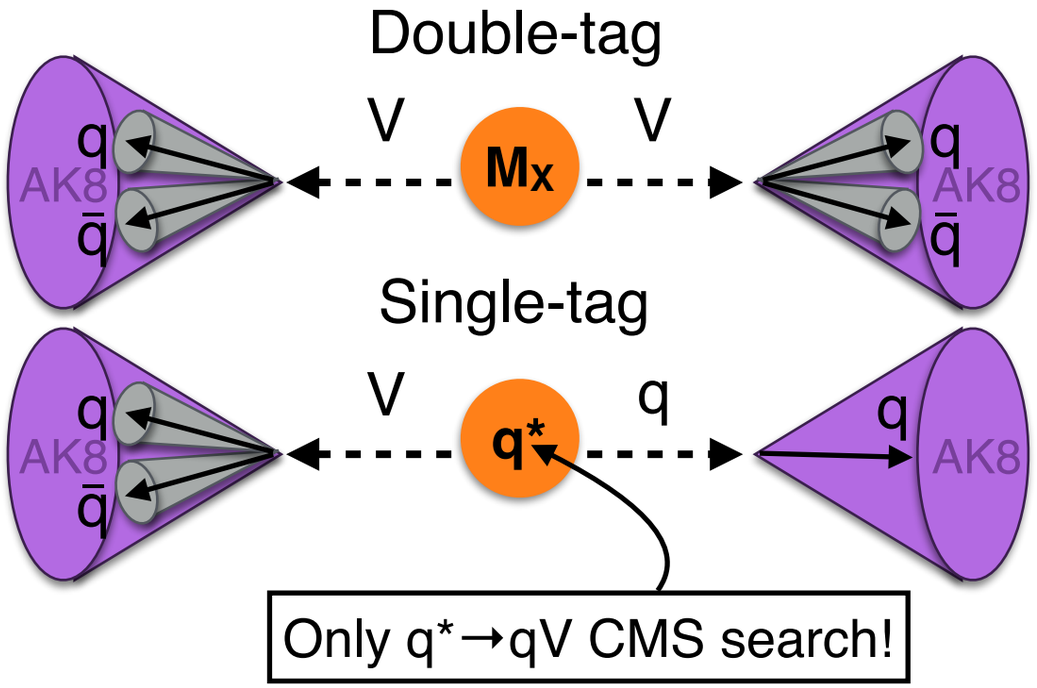
All-hadronic VV/qV

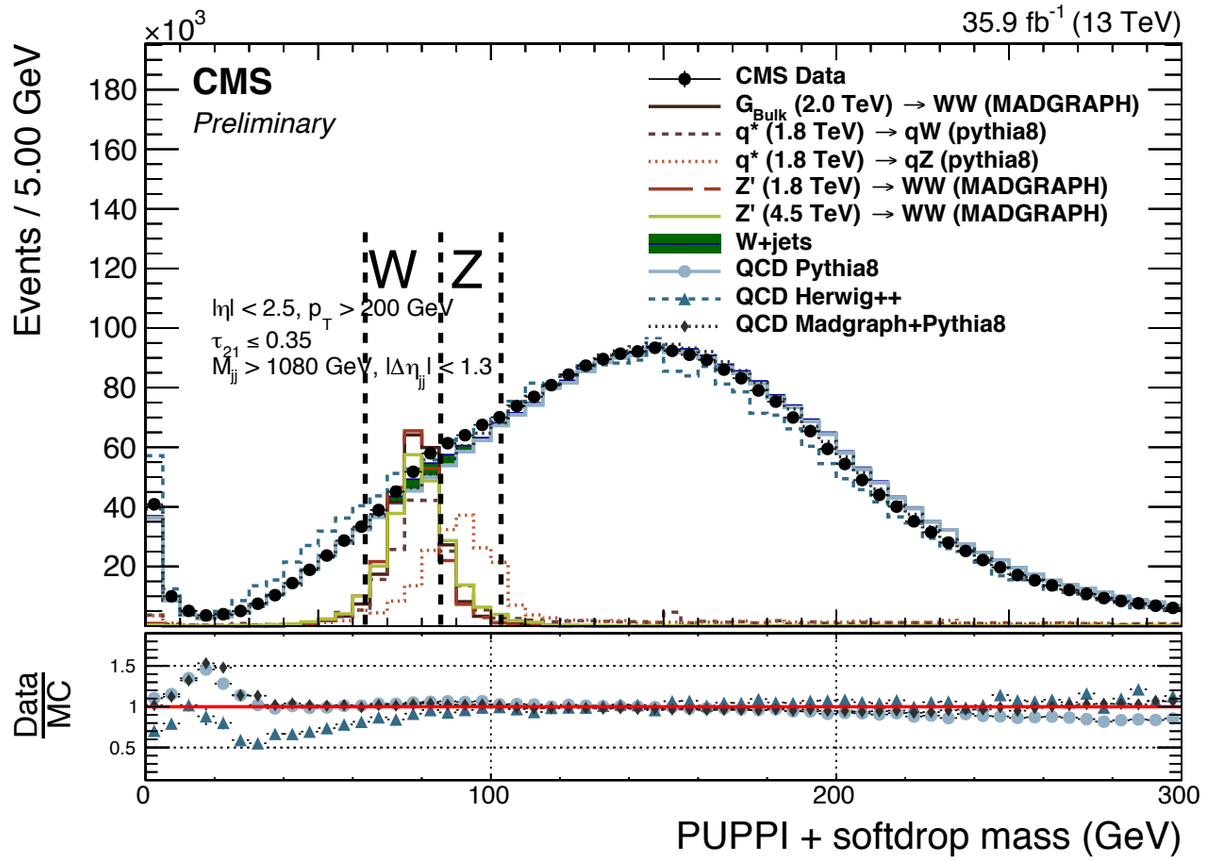
B2G-17-001

Categorisation (10 in total):

WW/WZ/ZZ and qW/qZ

τ_{21} high-purity / low-purity
(LP 20% improvement at high m_X)





All-hadronic VV/qV

B2G-17-001

Categorisation (10 in total):

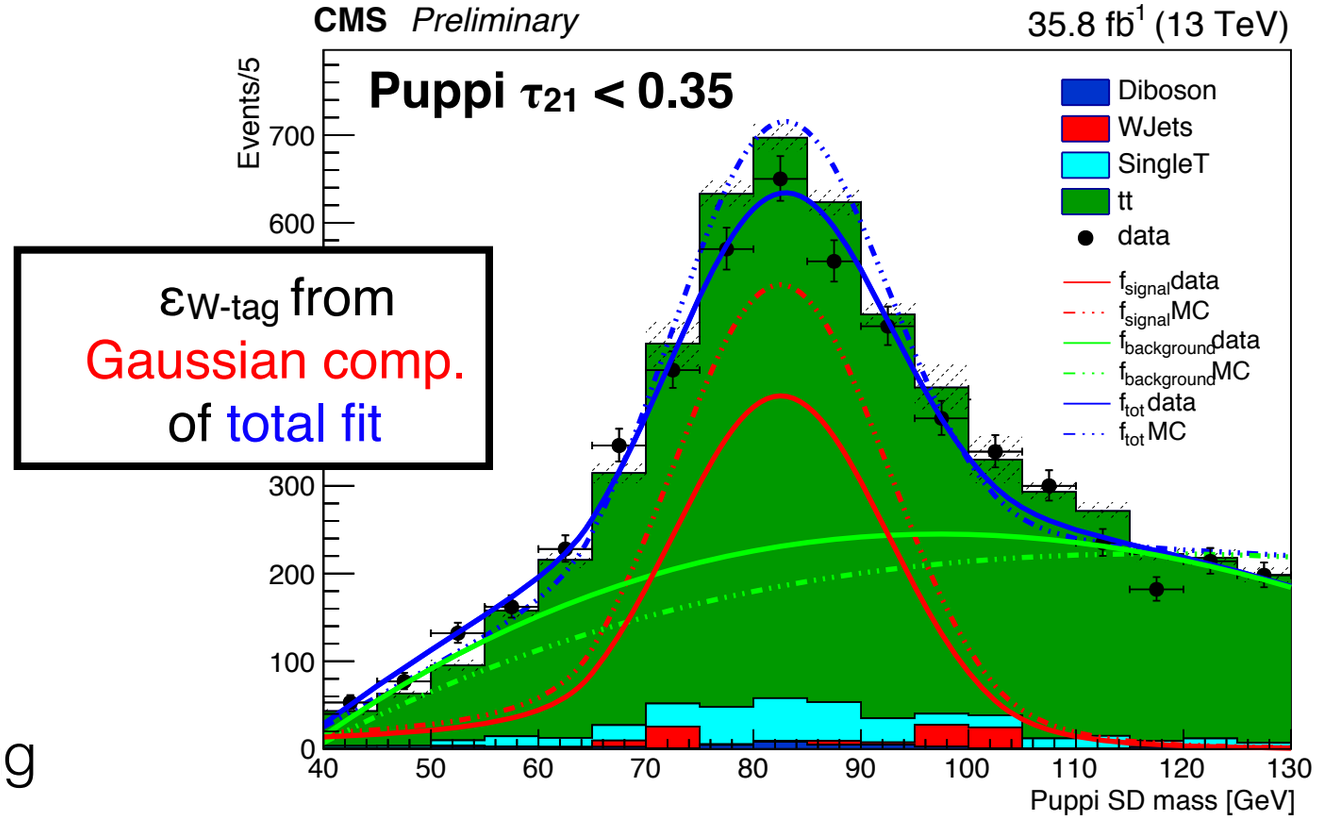
WW/WZ/ZZ and qW/qZ

τ_{21} high-purity / low-purity
(LP 20% improvement at high m_X)

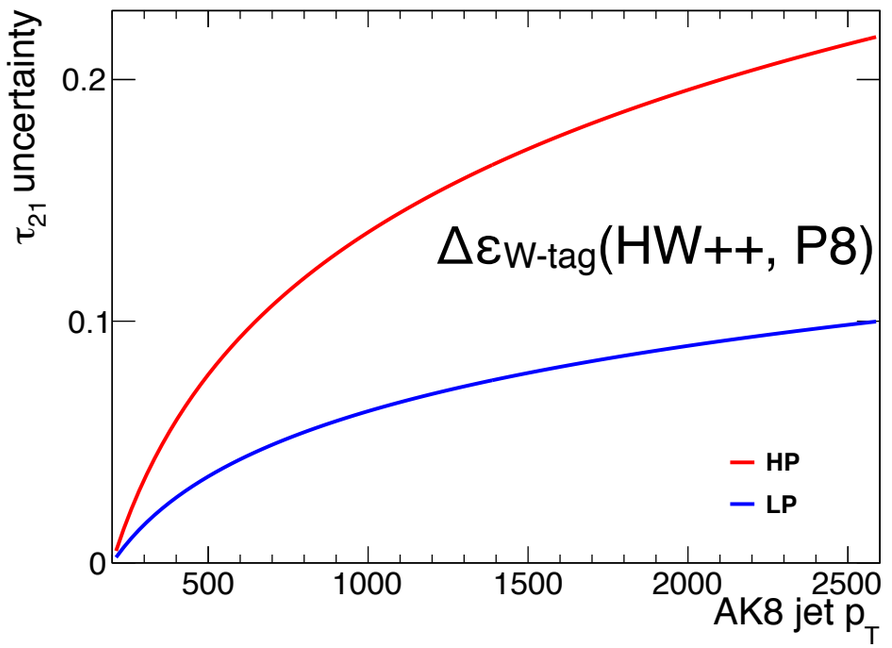
W-tag efficiency scalefactors, jet mass scale+res. estimated in $t\bar{t}$ (~200 GeV)

p_T-dependence is difference in tagging efficiency in HERWIG++ and Pythia8 signal MC

CMS DP-2017/026



εW-tag from Gaussian comp. of total fit



All-hadronic VV/qV

B2G-17-001

Background fit with parametric function in signal region

Fit function for background

$$\frac{dN}{dm_{jj}} = \frac{P_0}{(m_{jj}/\sqrt{s})^{P_2}} \text{ or } \frac{dN}{dm_{jj}} = \frac{P_0(1-m_{jj}/\sqrt{s})^{P_1}}{(m_{jj}/\sqrt{s})^{P_2}}$$

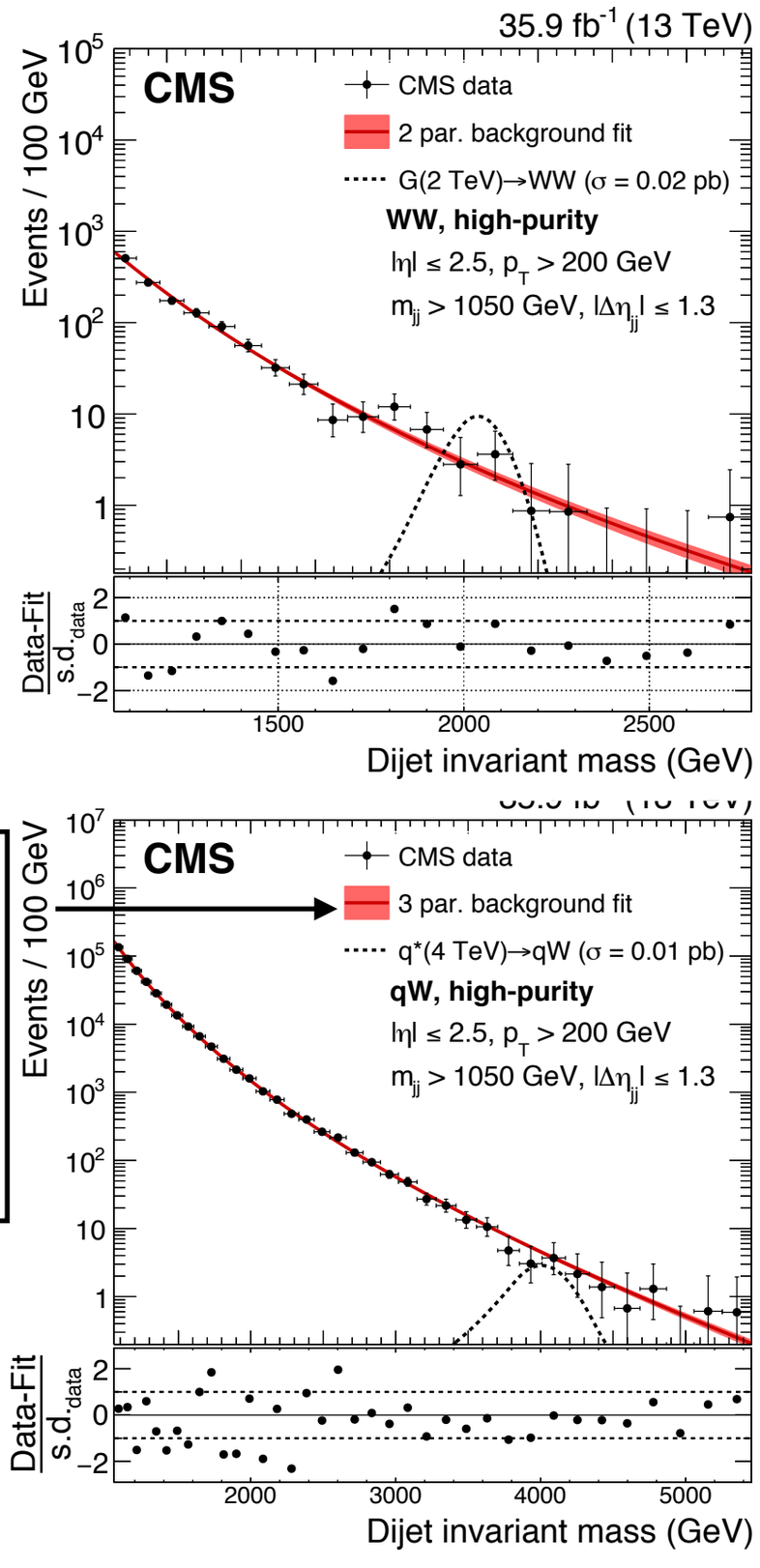
2 pars **3 pars**

… N pars

Determine number of parameters with F-test

check if quality of fit improves by 10% confidence level
if not, stick with current function

More parameters needed when statistics increase



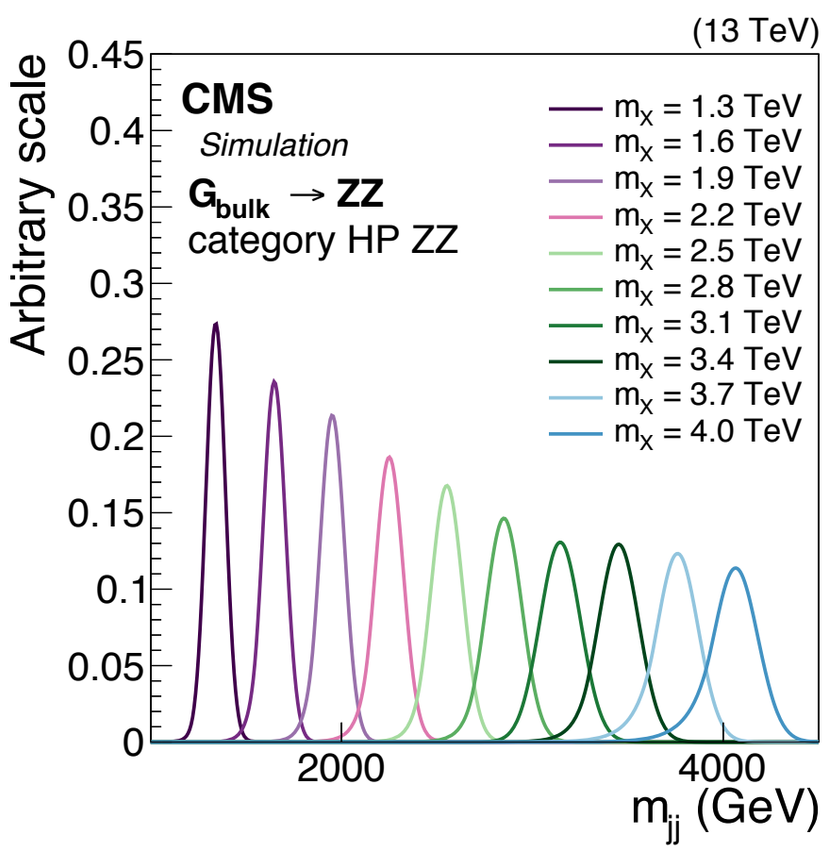
All-hadronic VV/qV

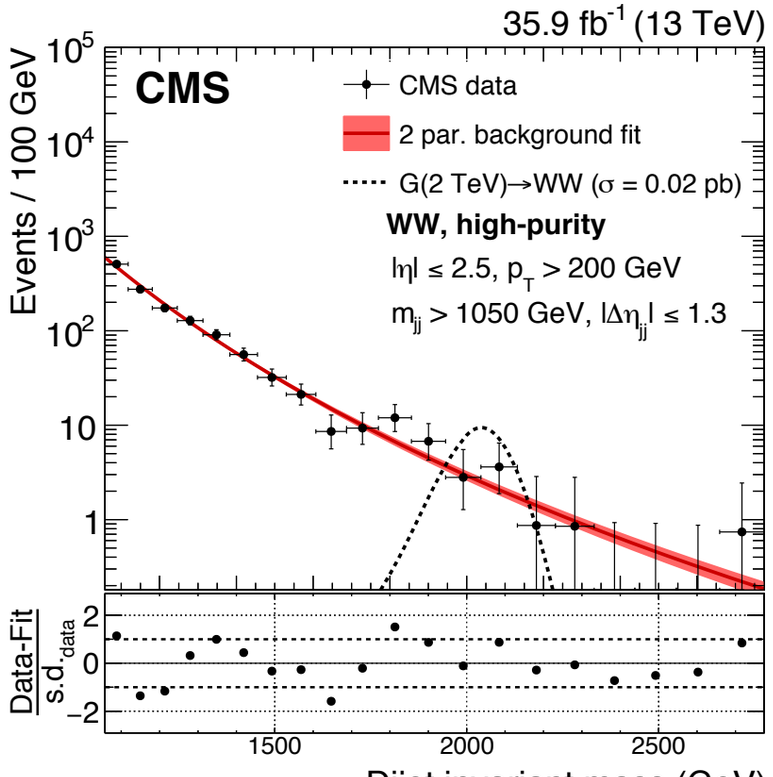
B2G-17-001

Background fit with parametric function in signal region

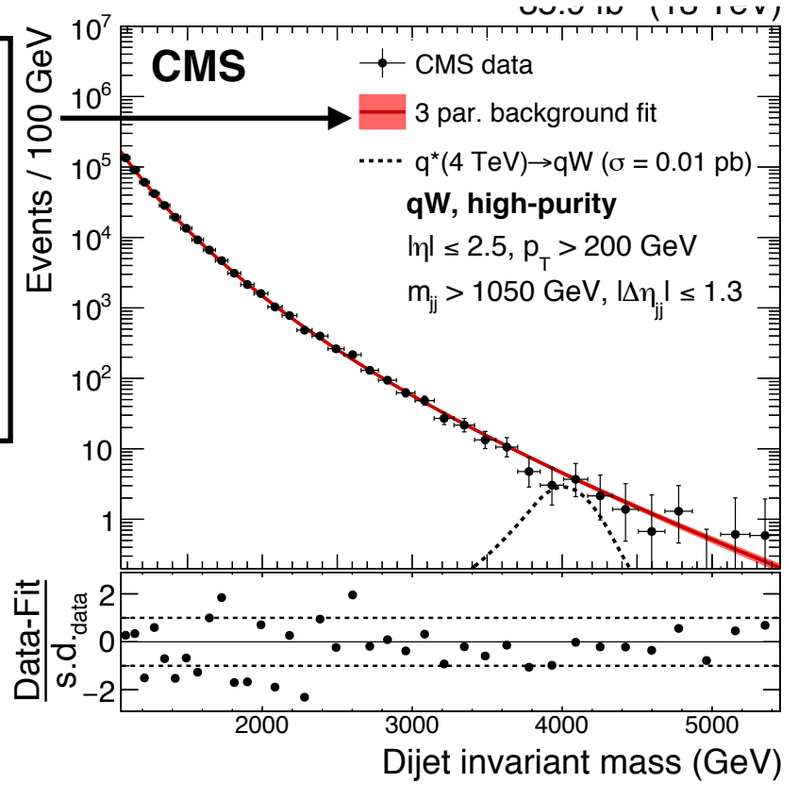
$$\frac{dN}{dm_{jj}} = \frac{P_0}{(m_{jj}/\sqrt{s})^{P_2}} \text{ or } \frac{dN}{dm_{jj}} = \frac{P_0(1-m_{jj}/\sqrt{s})^{P_1}}{(m_{jj}/\sqrt{s})^{P_2}}$$

Signal modeled with gaussian core with a power law tail to low masses





More parameters needed when statistics increase



All-hadronic VH

B2G-17-002

X→VH all-hadronic search very similar but two more categories for b-tagging (8 in total)

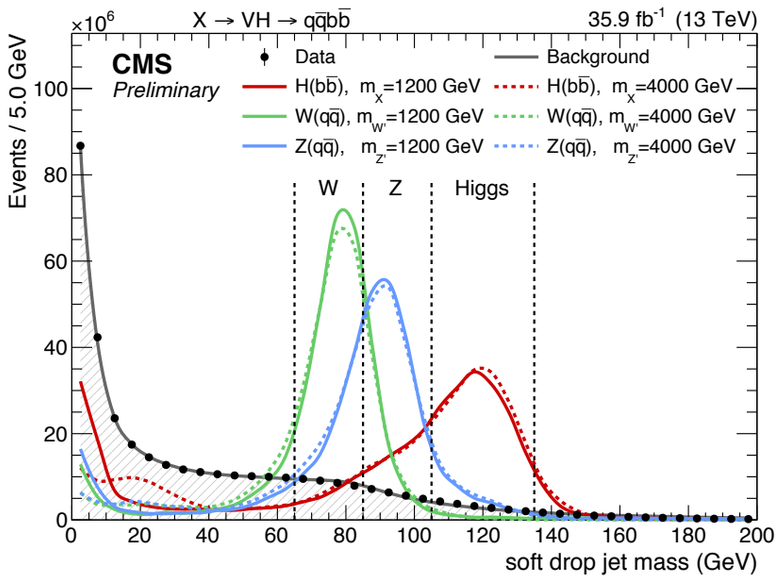
Background fit with parametric function in signal region

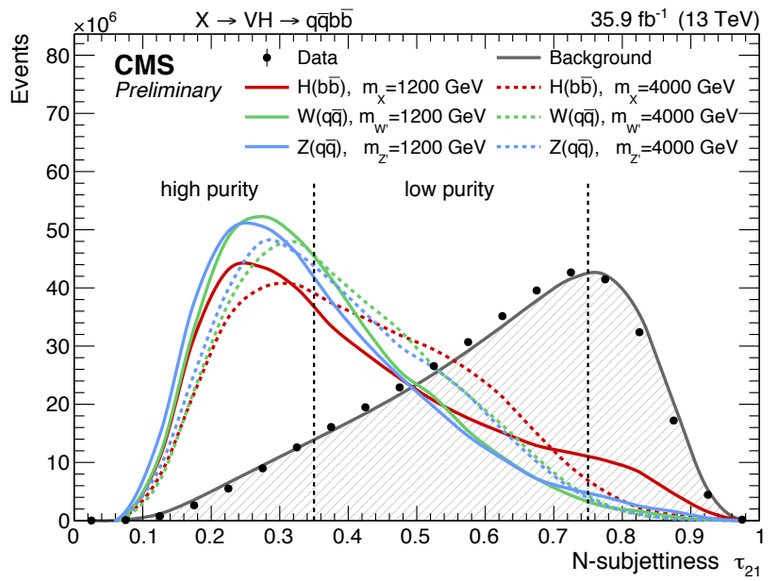
$$\frac{dN}{dm_{jj}} = \frac{P_0}{(m_{jj}/\sqrt{s})^{P_2}} \text{ or } \frac{dN}{dm_{jj}} = \frac{P_0(1-m_{jj}/\sqrt{s})^{P_1}}{(m_{jj}/\sqrt{s})^{P_2}}$$

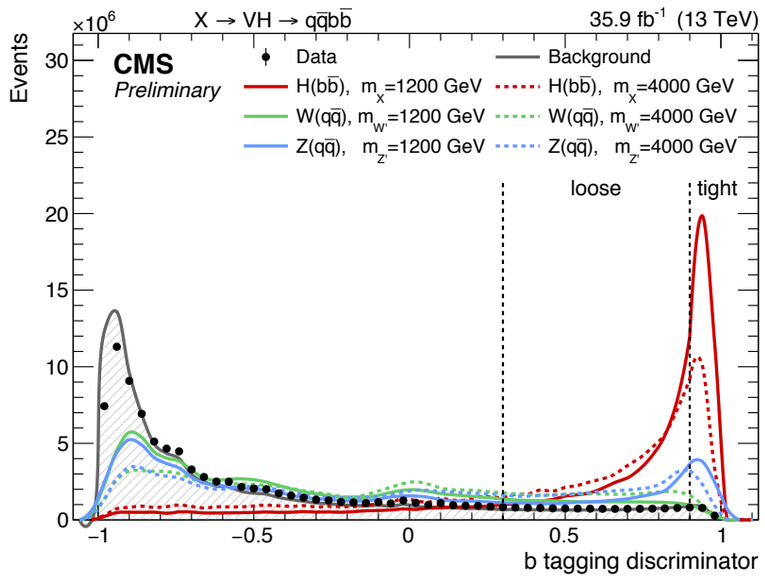
Signal modeled with gaussian core with a power law tail to low masses

B2G-17-001 [Phys. Rev. D 97, 072006 (2018)]

jet mass + τ_{21} + **b-tagging**







All-hadronic X→VV/VH: limits

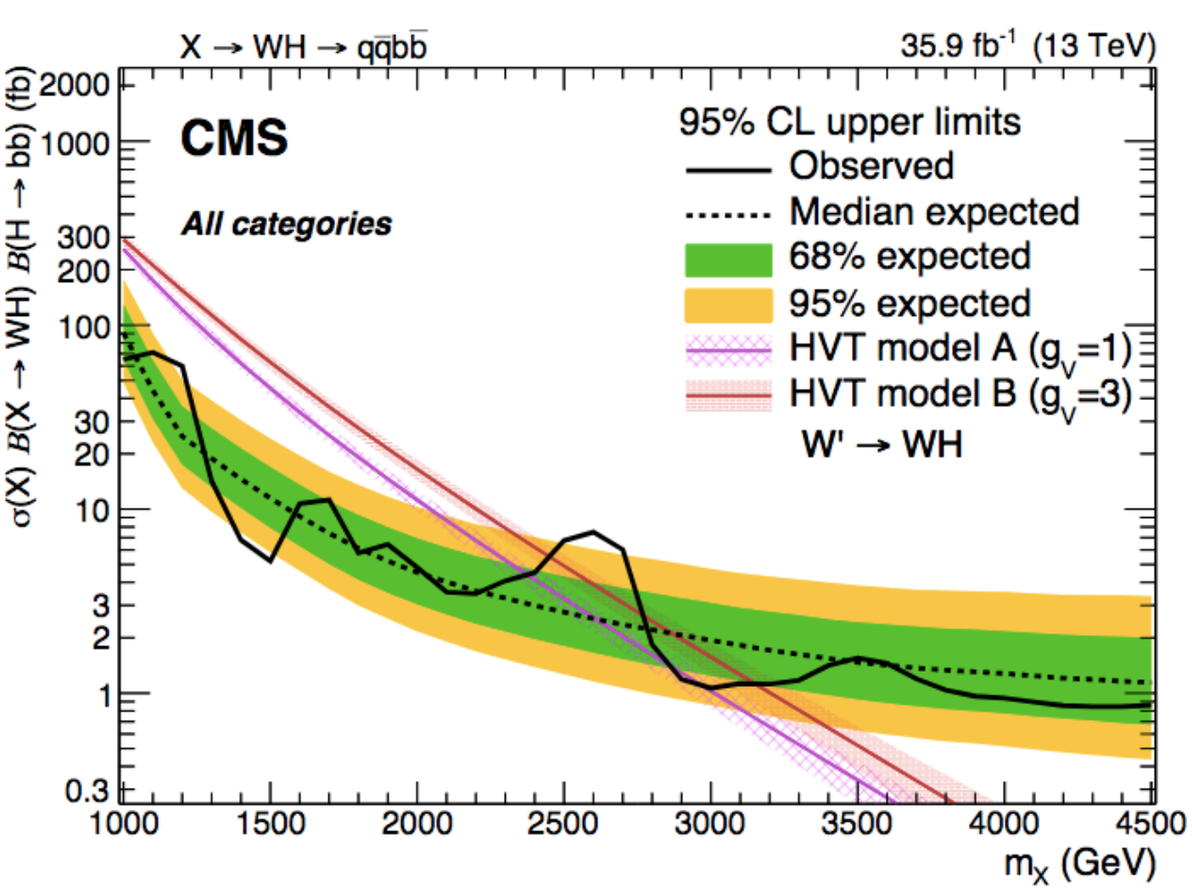
B2G-17-001

B2G-17-002

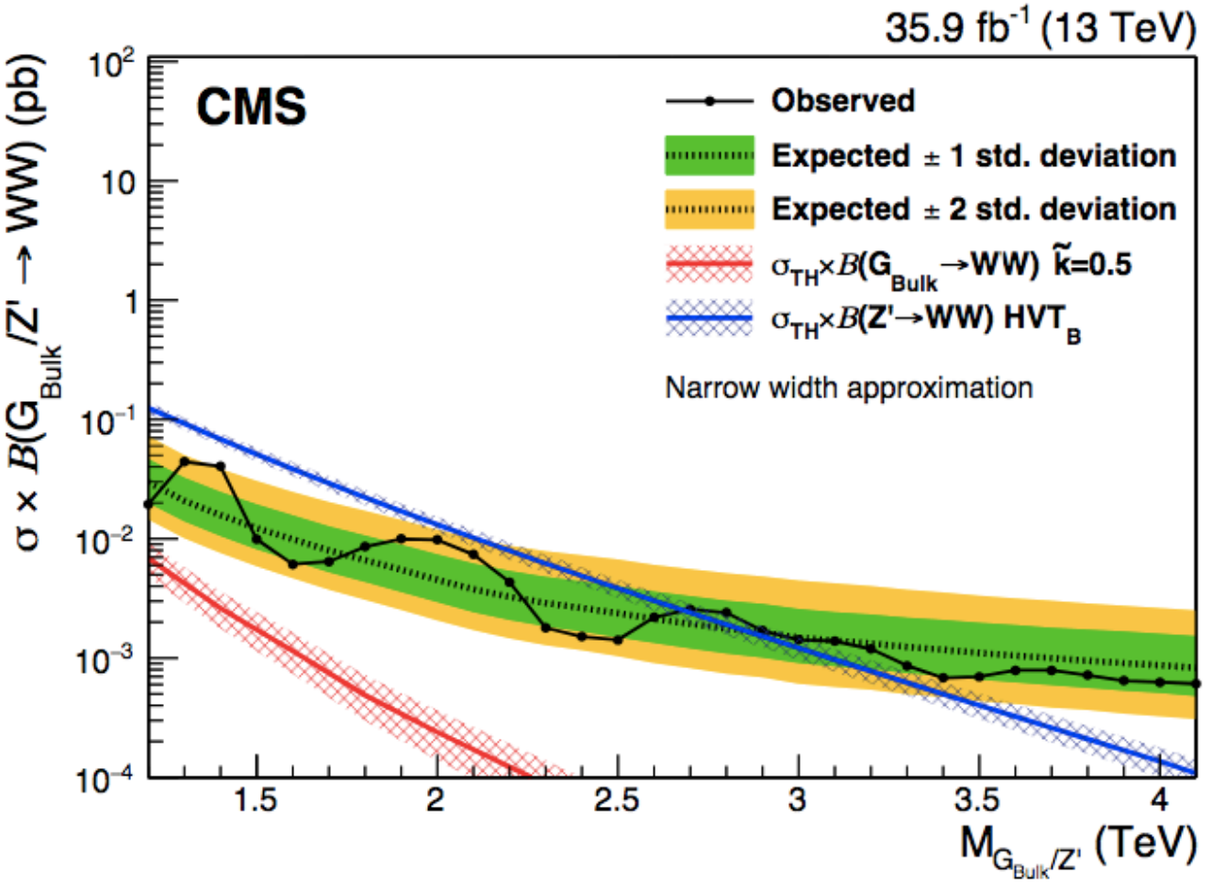
Results typically interpreted in the simplified **Heavy Vector Triplet (HVT) model**: generalizes explicit models predicting spin-1 W' and Z' resonances

Additionally, spin-2 interpretation for VV search

spin-1 W'→WH



spin-2 X→VV



All-hadronic X→VV/VH: limits

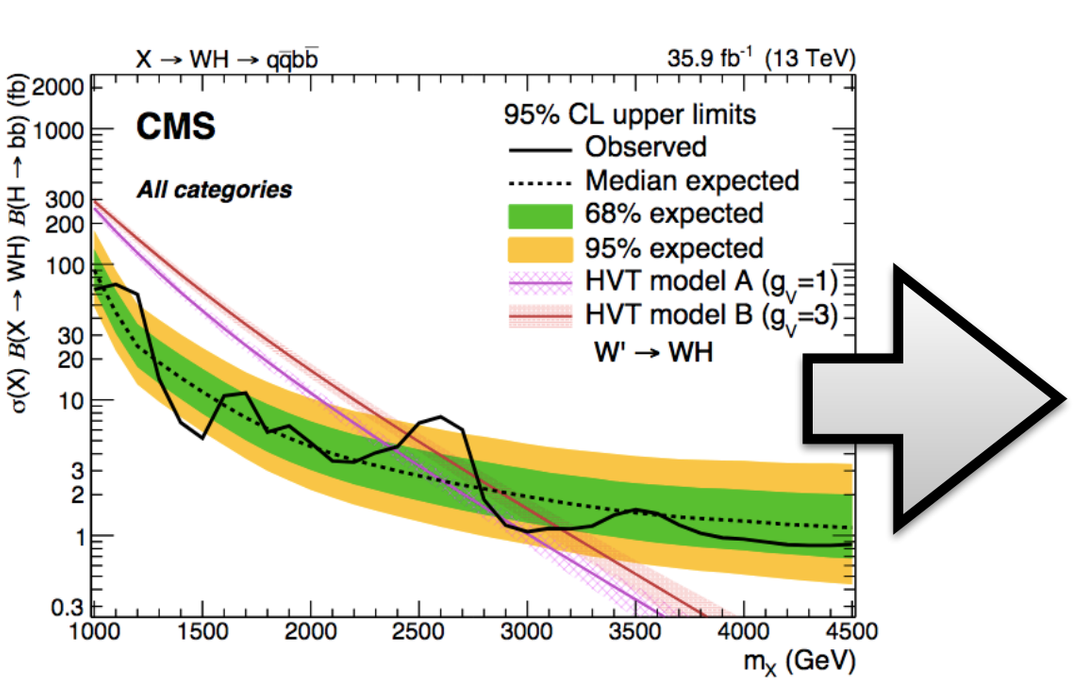
B2G-17-001

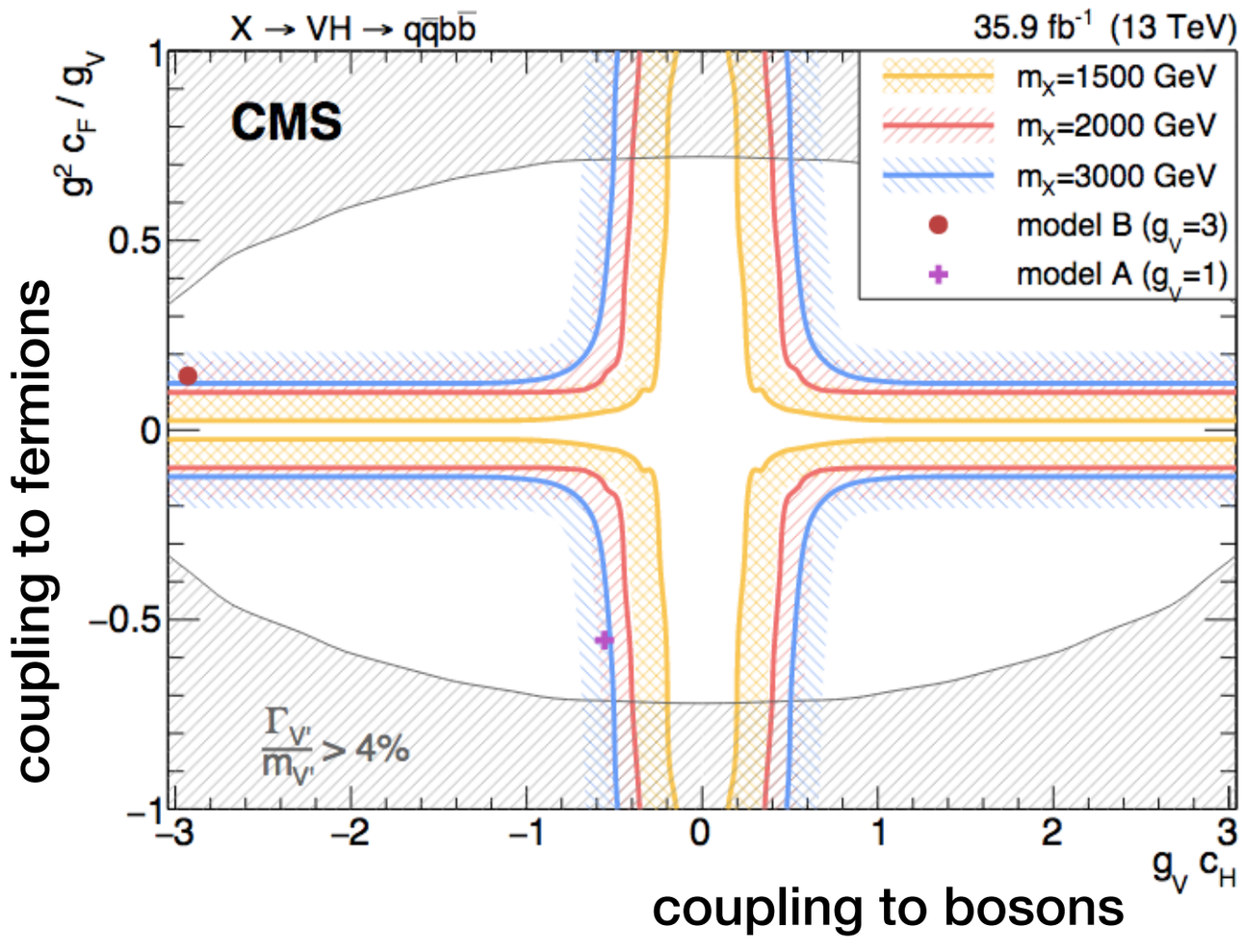
B2G-17-002

Results typically interpreted in the simplified **Heavy Vector Triplet (HVT) model**: generalizes explicit models predicting spin-1 W' and Z' resonances

Additionally, spin-2 interpretation for VV search

spin-1 W'→WH





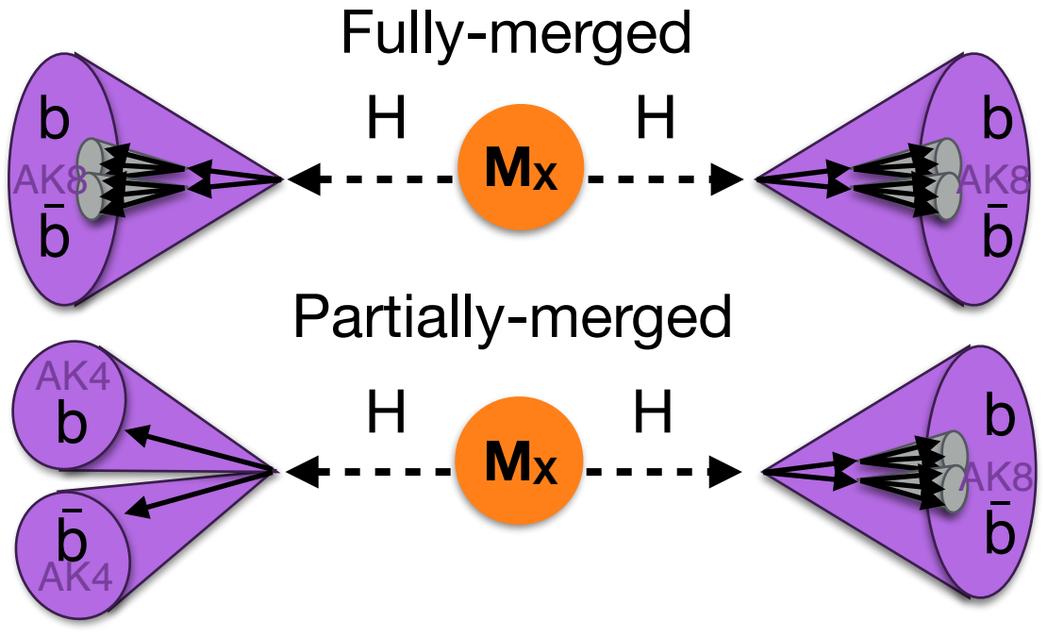
All-hadronic HH

B2G-17-019, B2G-16-026

Two complementary searches

fully-merged: tag two AK8 with double-b

partially-merged: tag one AK8 with double-b, tag two AK4 with CSV



All-hadronic HH

B2G-17-019, B2G-16-026

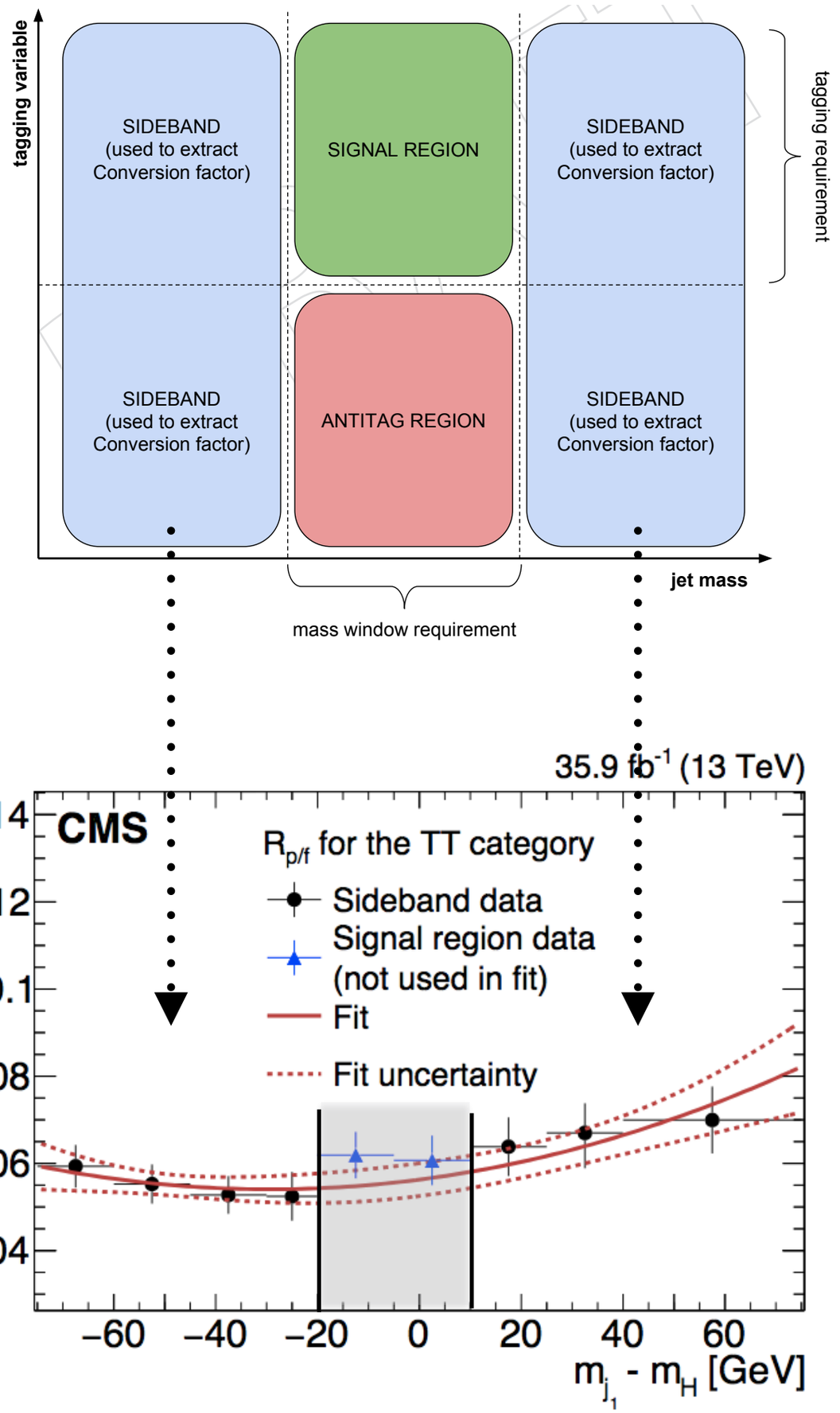
Two complementary searches

fully-merged: tag two AK8 with double-b

partially-merged: tag one AK8 with double-b, tag two AK4 with CSV

Background normalisation: mass sidebands

measure pass-fail ratio in data for double-b tagger and extrapolate to jet mass signal region



All-hadronic HH

B2G-17-019, B2G-16-026

Two complementary searches

fully-merged: tag two AK8 with double-b

partially-merged: tag one AK8 with double-b, tag two AK4 with CSV

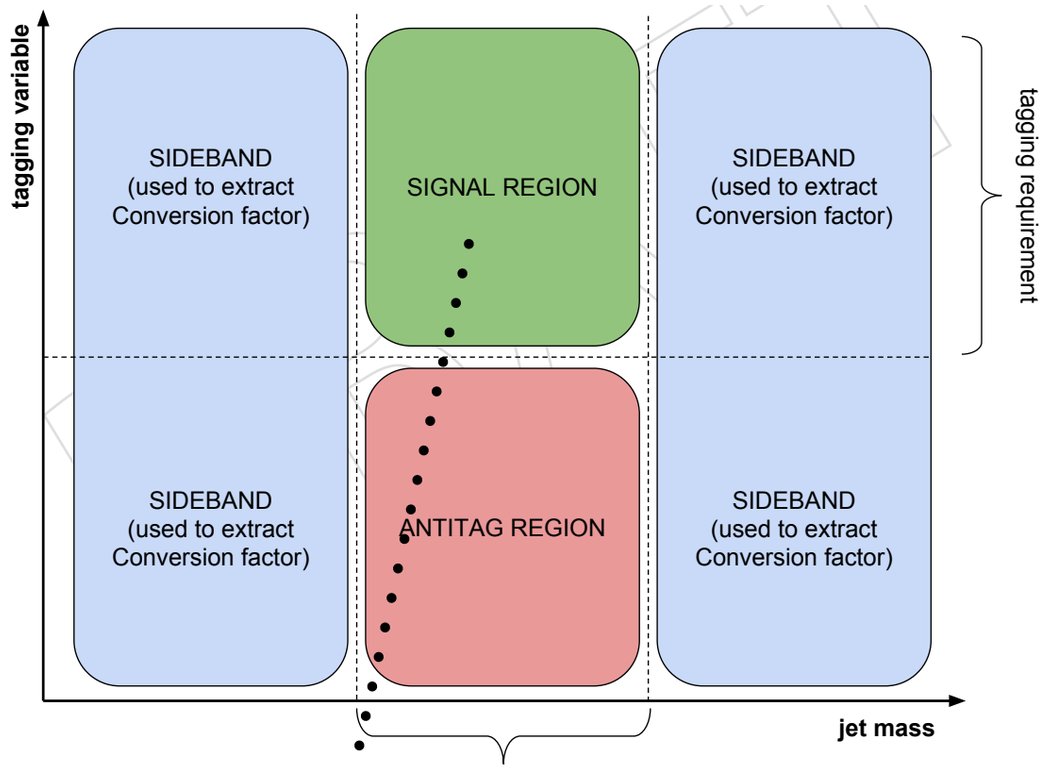
Background normalisation: mass sidebands

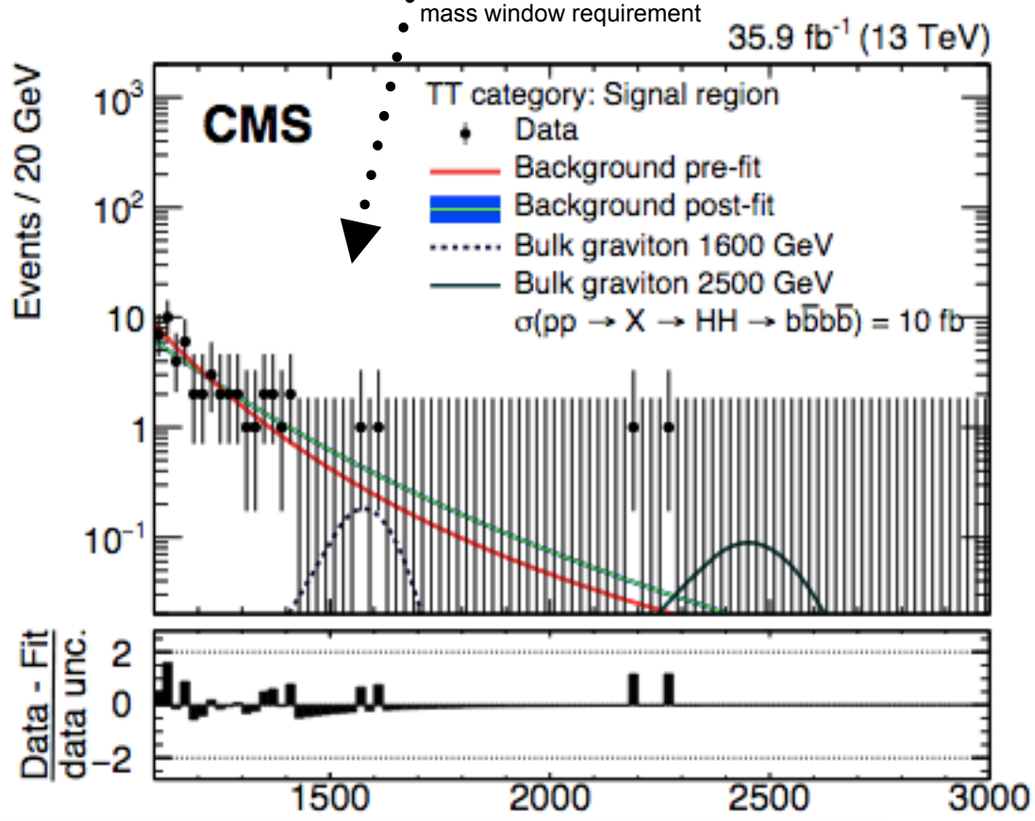
measure pass-fail ratio $R_{p/f}$ in data for double-b tagger and extrapolate to jet mass signal region

Background shape: anti b-tag region

simultaneous fit of $m_{jj,red}$ spectrum in antitag and signal region with smoothly falling parametric function

constrain normalization from $R_{p/f}$



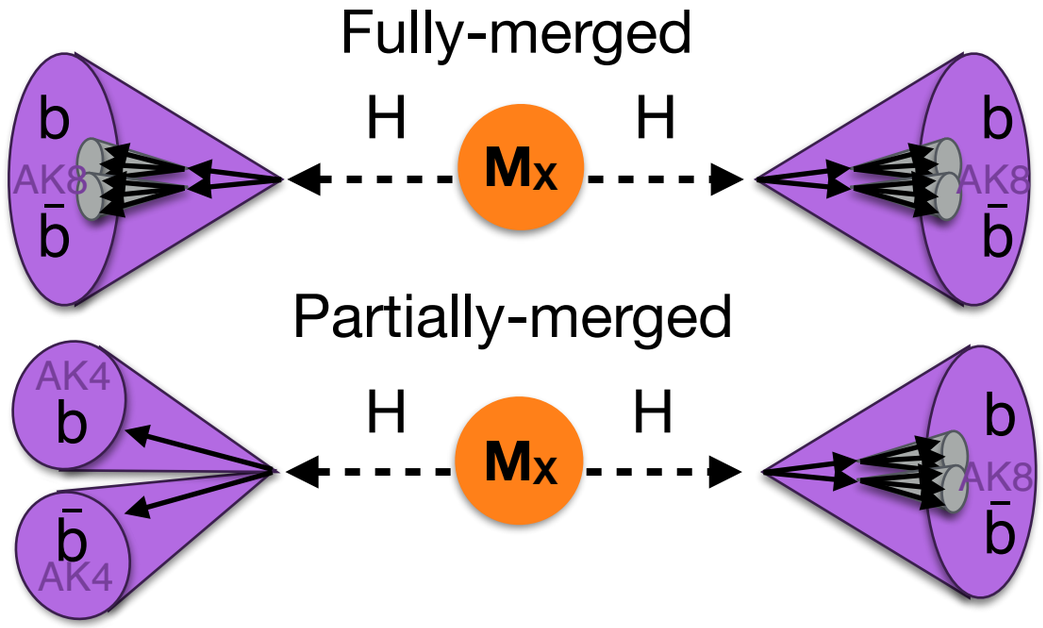


$m_{jj,red} \equiv m_{jj} - (m_J - m_H) - (m_{jj}(j_1, j_2) - m_H)$

8-10% improvement on HH mass resolution

All-hadronic HH

B2G-17-019, B2G-16-026



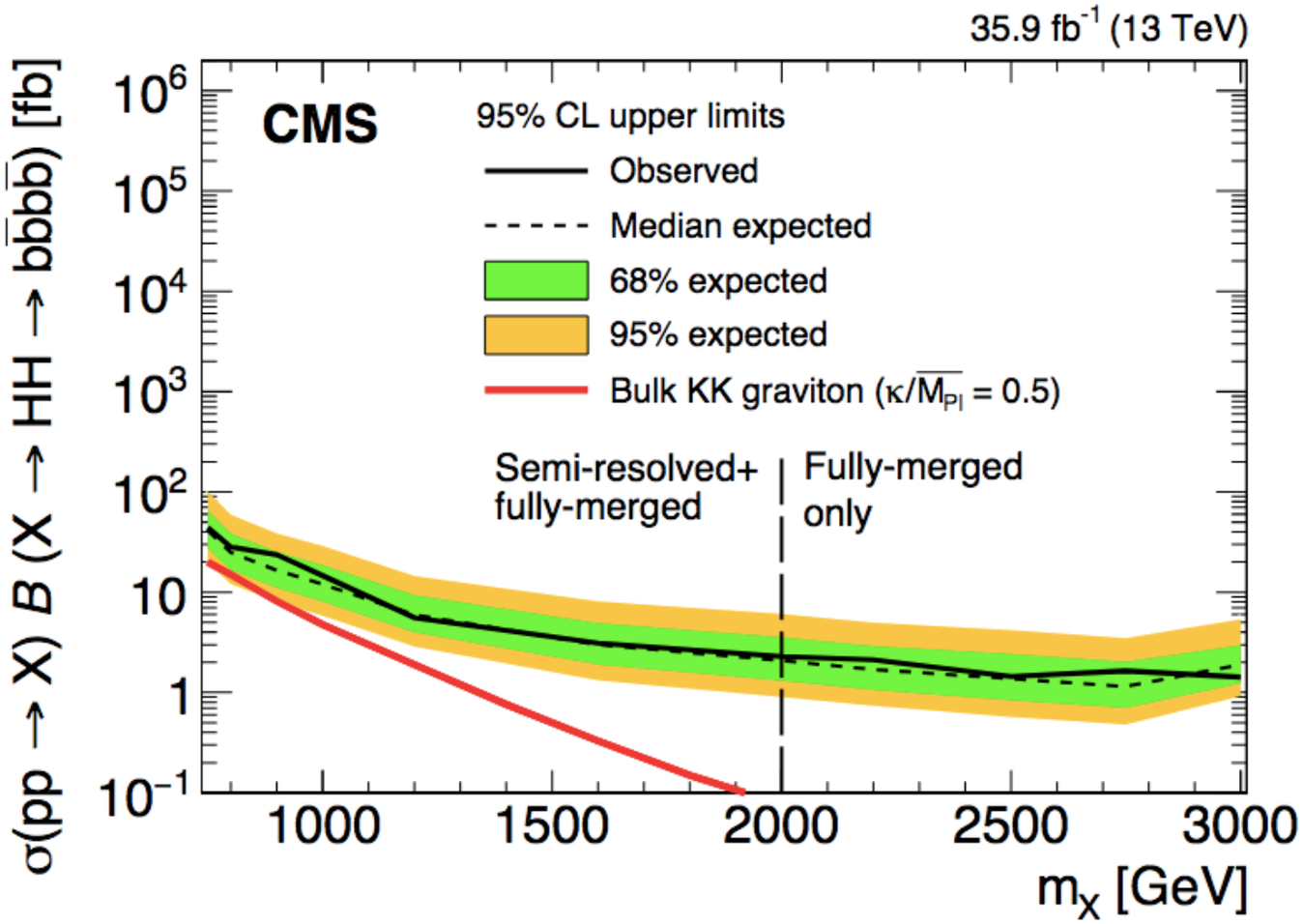
Two complementary searches

fully-merged: tag two AK8 with double-b

partially-merged: tag one AK8 with double-b, tag two AK4 with CSV

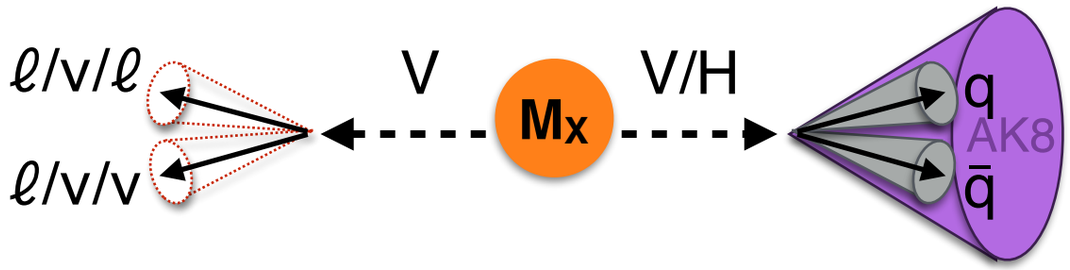
Adding partially-merged result improves limits by up to 55% !

resonant spin-0/2 interpretation



Semi-leptonic final states

X→VV or VH where one V decays leptonically (ℓℓ/ℓν/νν)

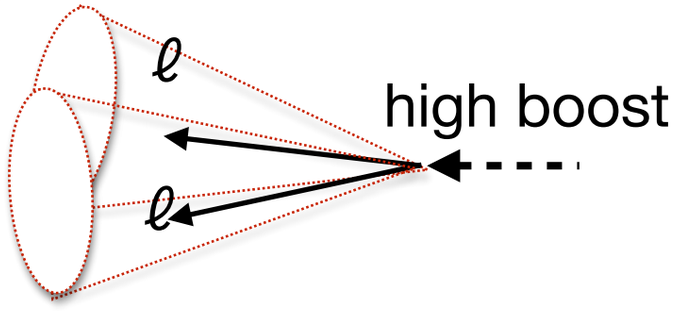


The pros:

can trigger on lepton, lower thresholds, while retaining good signal efficiency through hadronic leg

less background (most sensitive at low m_X) and background well-modelled in simulation (unlike QCD)

Isolation cone
ΔR < 0.3



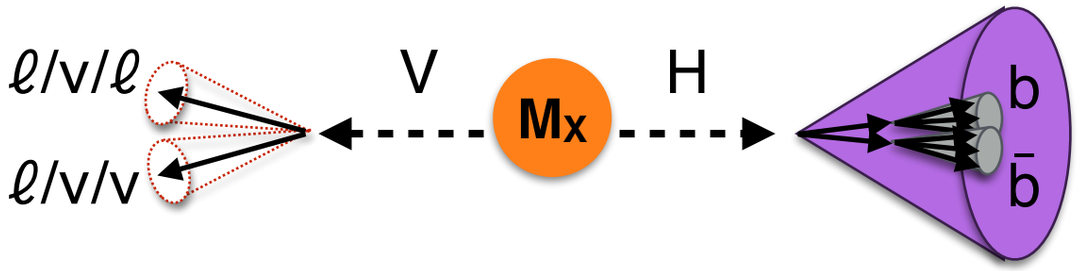
With standard reconstruction, leptons would be rejected due to non-isolation

The cons:

special care in dilepton channels to identify two isolated close leptons

VH semi-leptonic

B2G-17-004



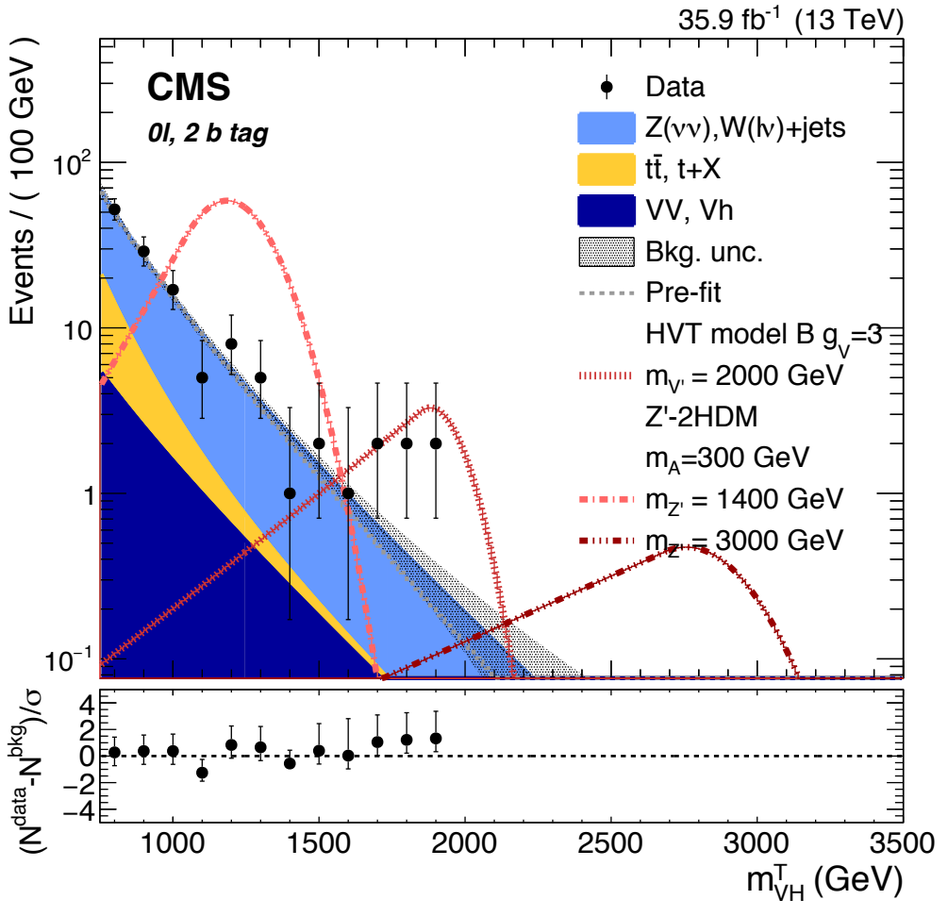
3 final states: H→bb̄ with V→ℓℓ/ℓν/νν (orthogonal 0,1,2 lepton categories)

Observable: m_{VH}^{T} for νν, m_{VH} for ℓℓ/ℓν

2 b-tag categories based on number of subjet b-tags:

2 b-tagged subjets: dominate at low m_X
→ ε_S 29-19% (degrade with m_X)
→ ε_B ~0.5%

1 b-tagged subjet: dominate at high m_X
→ ε_S 13-24% → ε_B ~3%



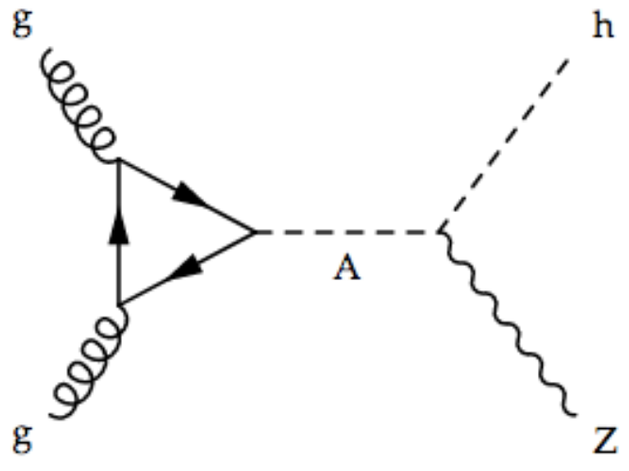
Semi-leptonic X→VH: limits

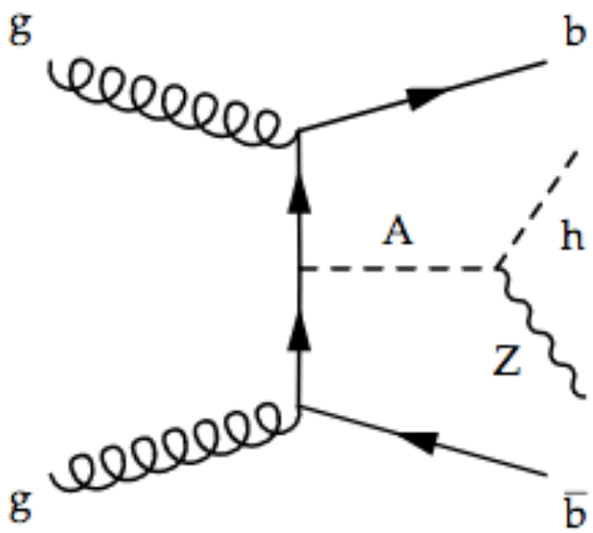
B2G-17-004

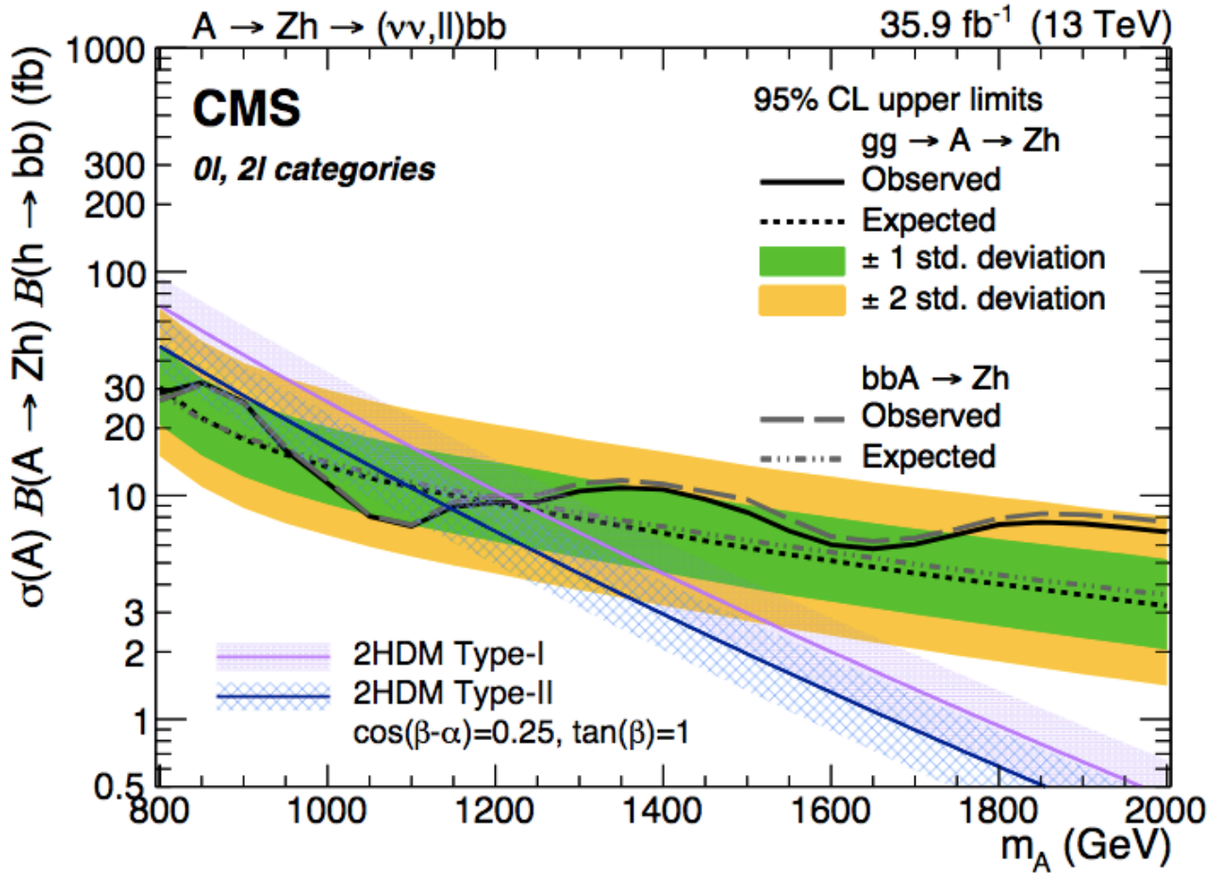
Results typically interpreted HVT model for spin-1 W' and Z' resonances

! For the first time, interpretation in the context of **spin-0 two Higgs doublet model (2HDM)**

Type-I and Type-II: predict a neutral and CP-odd boson A







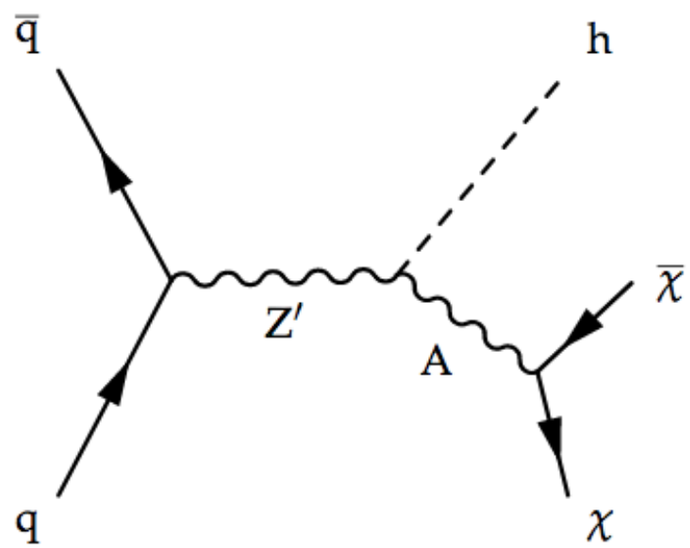
Semi-leptonic X→VH: limits

B2G-17-004

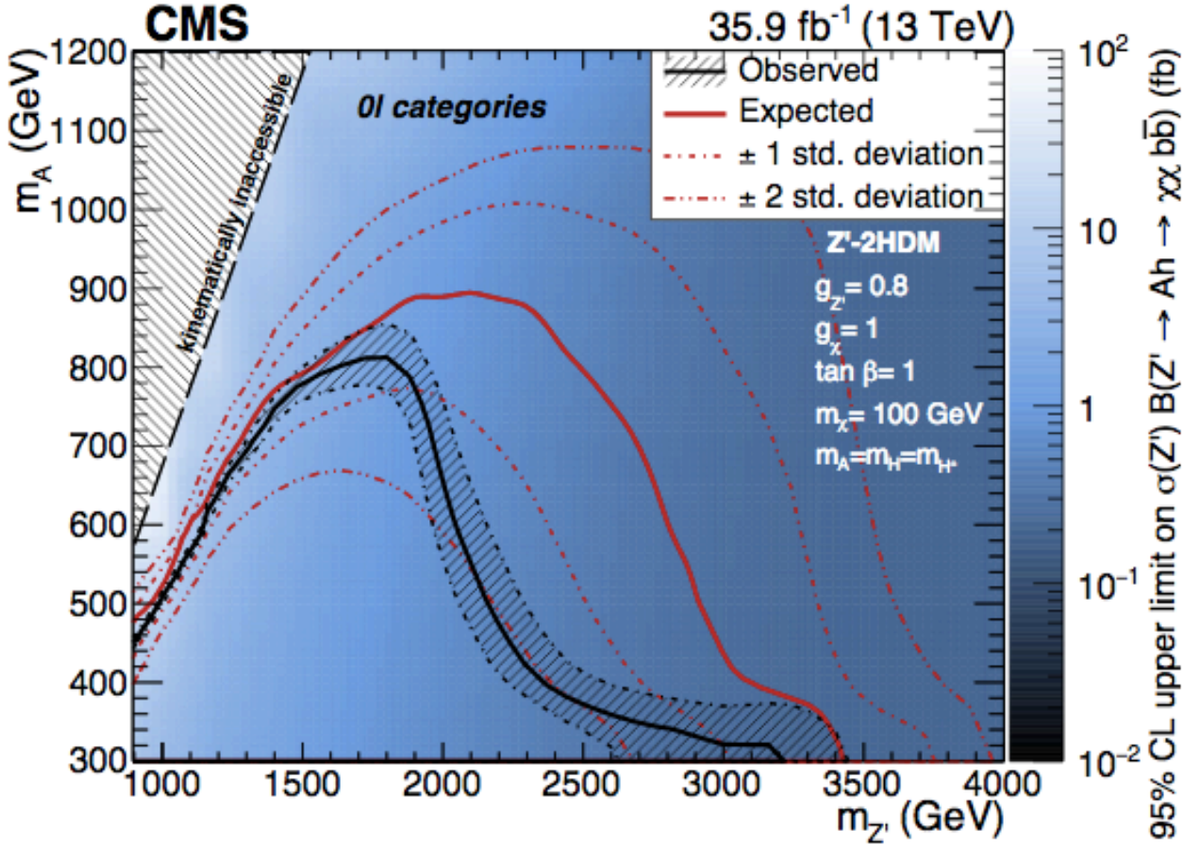
Results typically interpreted HVT model for spin-1 W' and Z' resonances

! For the first time, interpretation in the context of **spin-0 two Higgs doublet models (2HDM)**

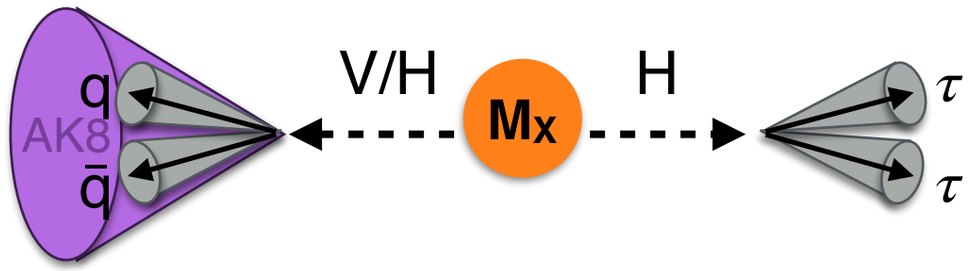
Z'-2HDM: introduces Z' particle + candidate for dark matter coupling to A boson



Most stringent limits to date on Z'-2HDM!



Boosted τ final states

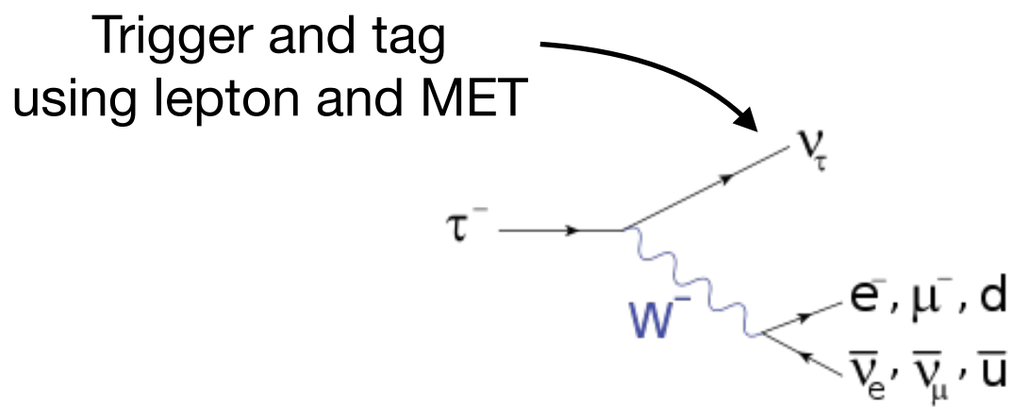


$X \rightarrow VH \rightarrow q\bar{q}\tau\tau$ or $X \rightarrow HH \rightarrow b\bar{b}\tau\tau$

The pros:

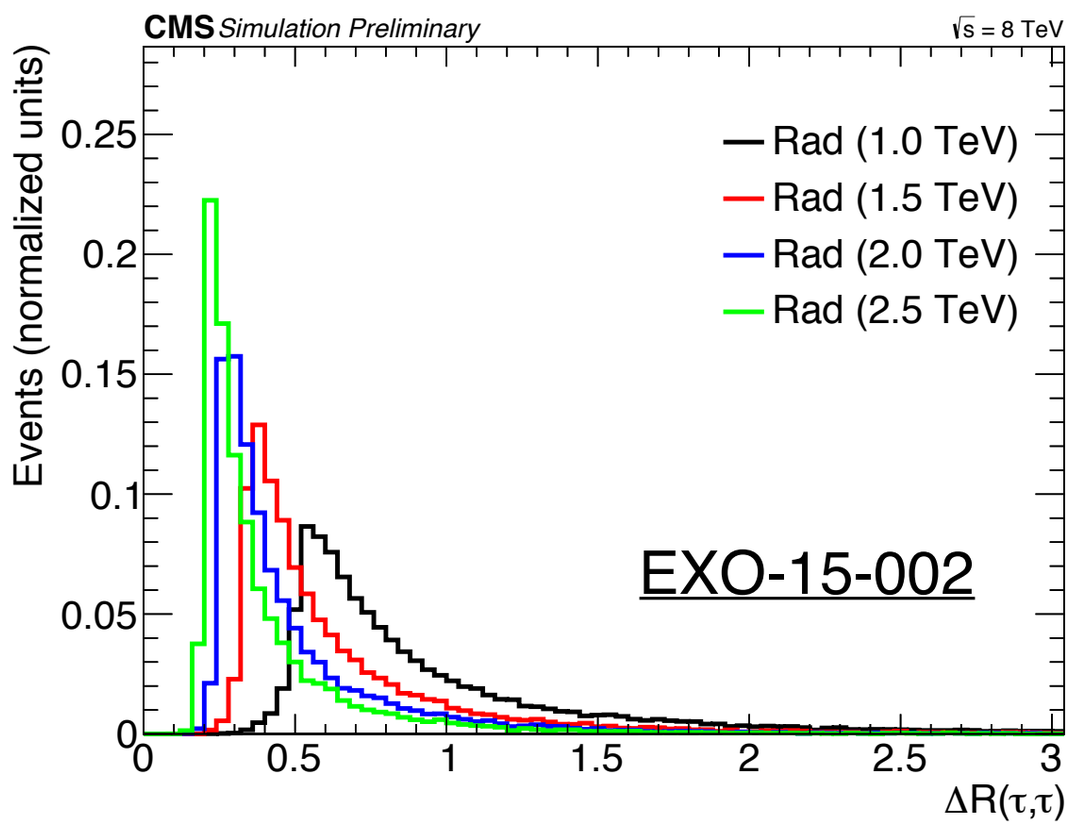
high branching fraction

with one τ decaying leptonically, low background and possibility to trigger on lepton+MET



The cons:

extremely challenging to reconstruct $H \rightarrow \tau\tau$ as τ decay products overlap



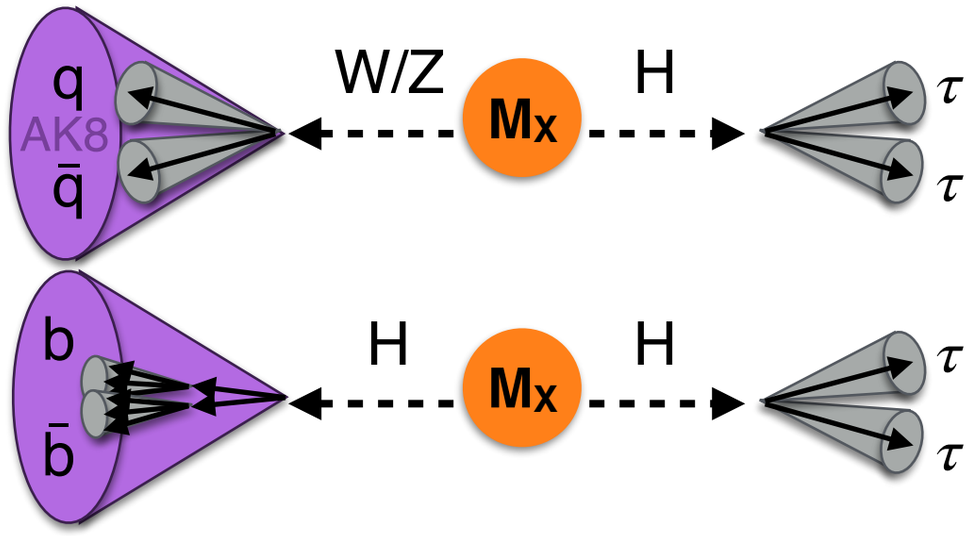
VH/HH to $\tau\tau$qq /$\tau\tau$bb

B2G-17-006

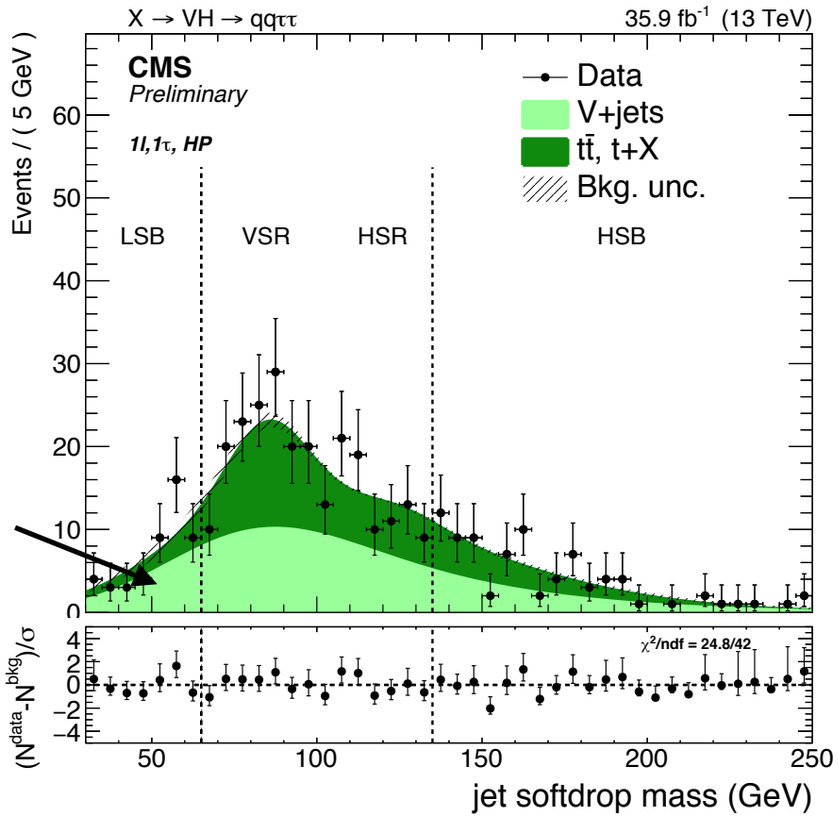
3 orthogonal searches: WH/ZH/HH

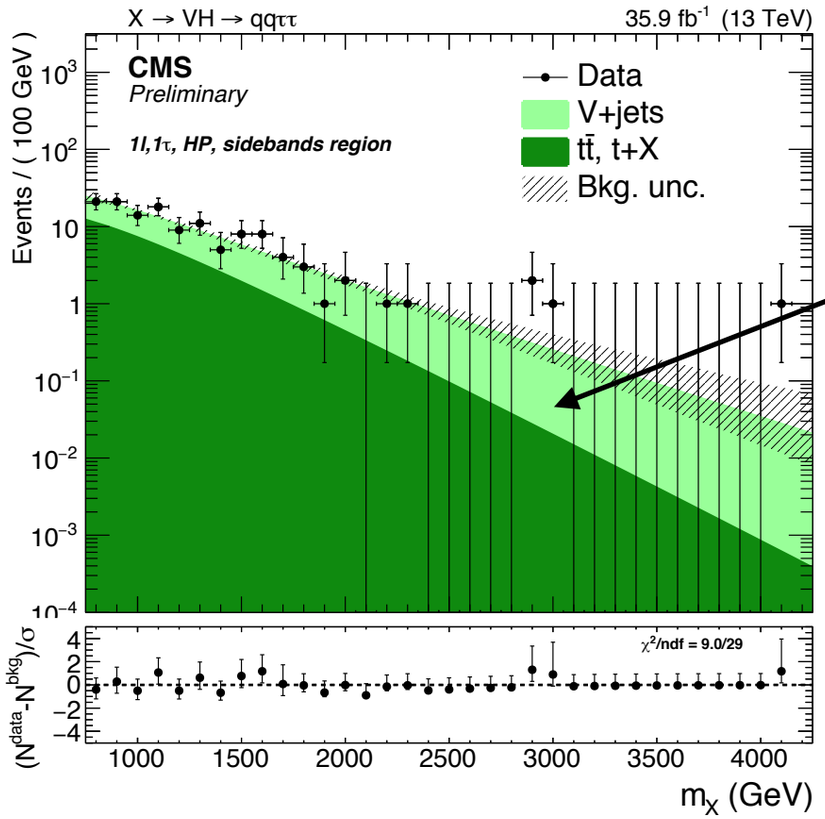
Combine subjet b-tagging, boosted V reconstruction and boosted τ tagging

both fully-hadronic and semi-leptonic $\tau\tau$ final states using boosted reconstruction



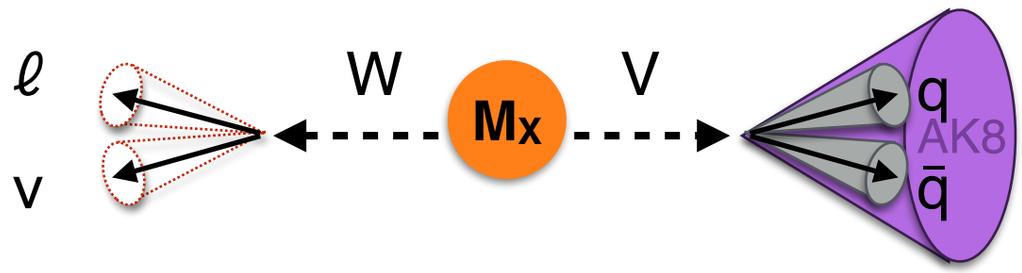
normalisation from mass sidebands





shape from m_X sideband, extrapolated to signal region via transfer function

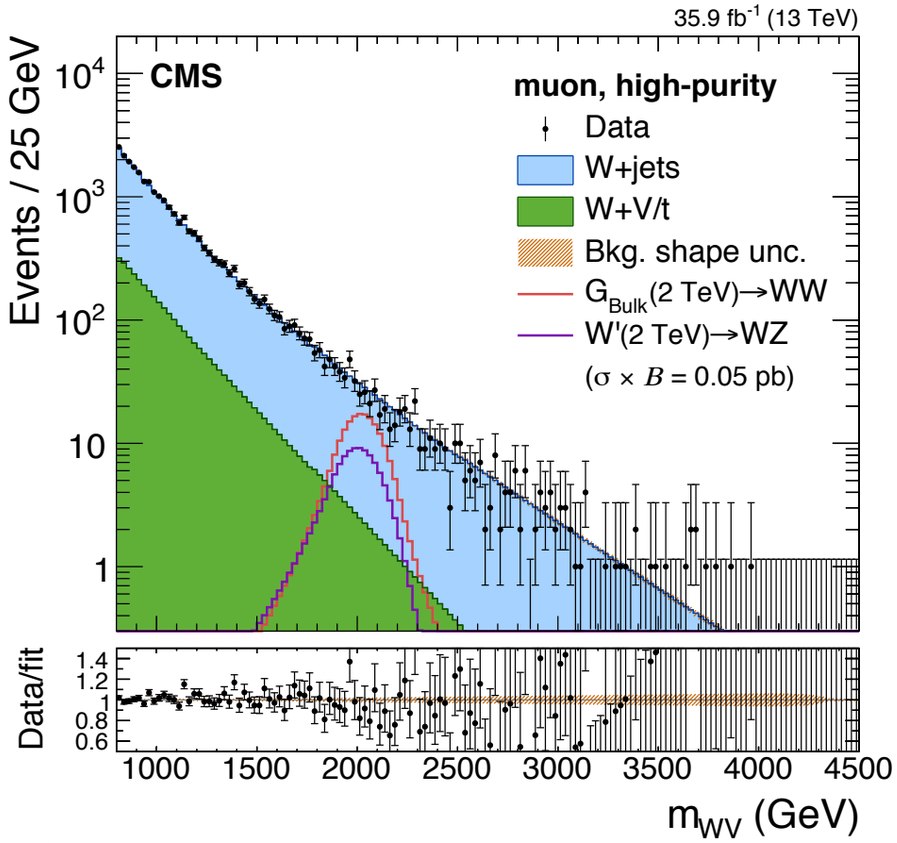
Going multidimensional



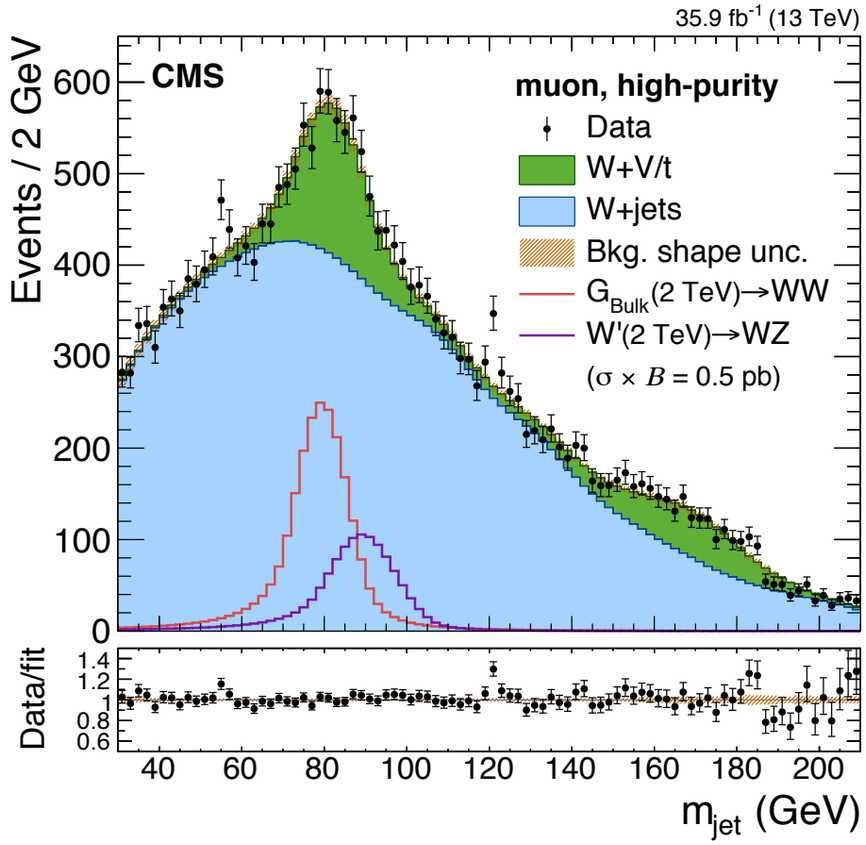
Take advantage of signal peaking in both jet mass and invariant mass and search for X→VW in $M_{\ell\nu,jet}$-m_{jet} plane

$$P_{\mathrm{W+jets}}(m_{\mathrm{WV}}, m_{\mathrm{jet}}) = P_{\mathrm{WV}}(m_{\mathrm{WV}}|m_{\mathrm{jet}}, \theta_1)\, P_j(m_{\mathrm{jet}}|\theta_2)$$

diboson invariant mass $M_{\ell\nu,jet}$

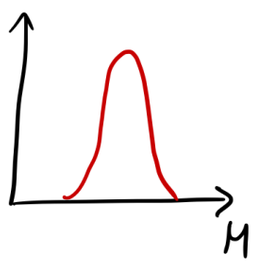


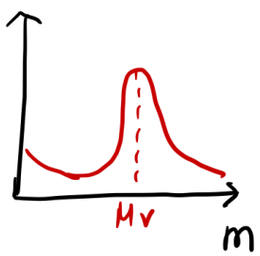
jet mass m_{jet}



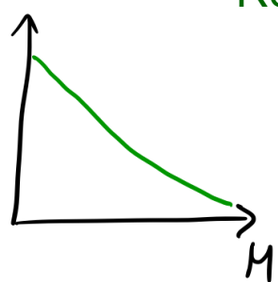
B2G-16-019

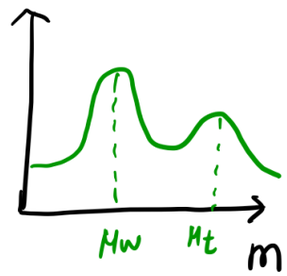
Signal



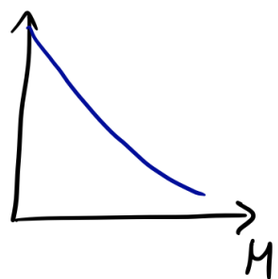


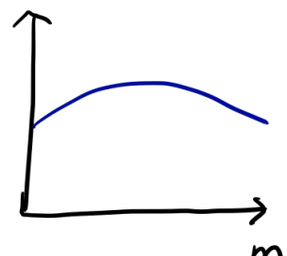
Resonant





Non-resonant

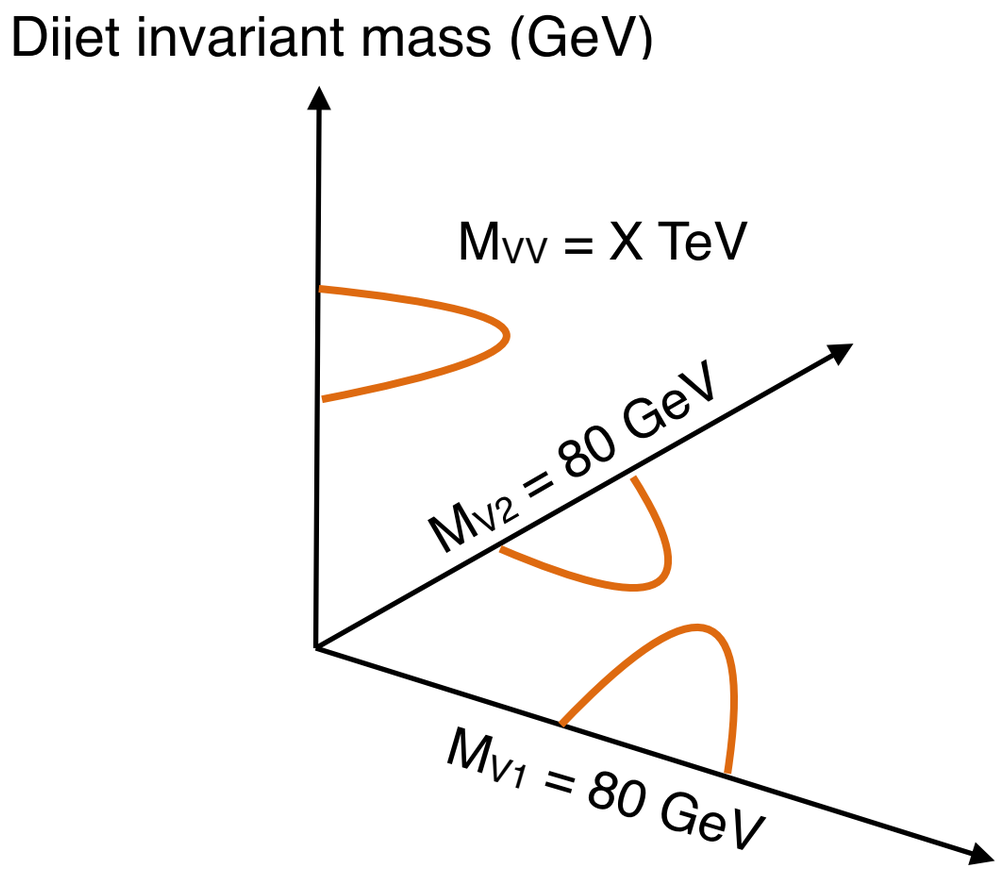




Outlook: 2D → 3D

For all-hadronic, extend this to 3D scan of M_{V1}-M_{V2} - M_{VV}

use full jet mass line shape, can incorporate all VV/VH searches in one common framework



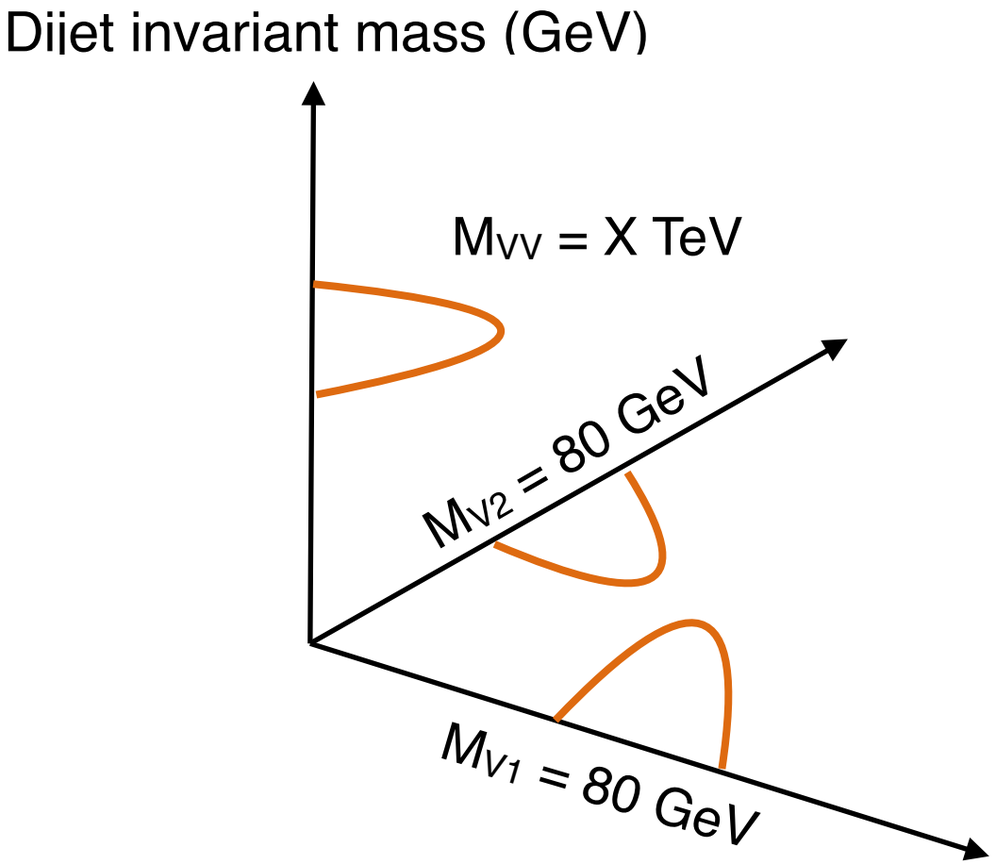
Outlook: 2D → 3D

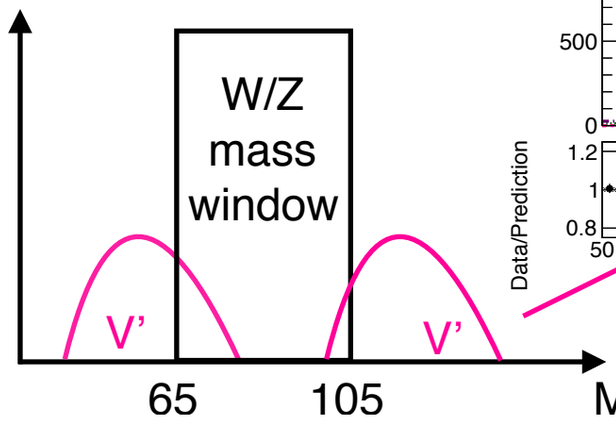
For all-hadronic, extend this to 3D scan of M_{V1}-M_{V2} - M_{VV}

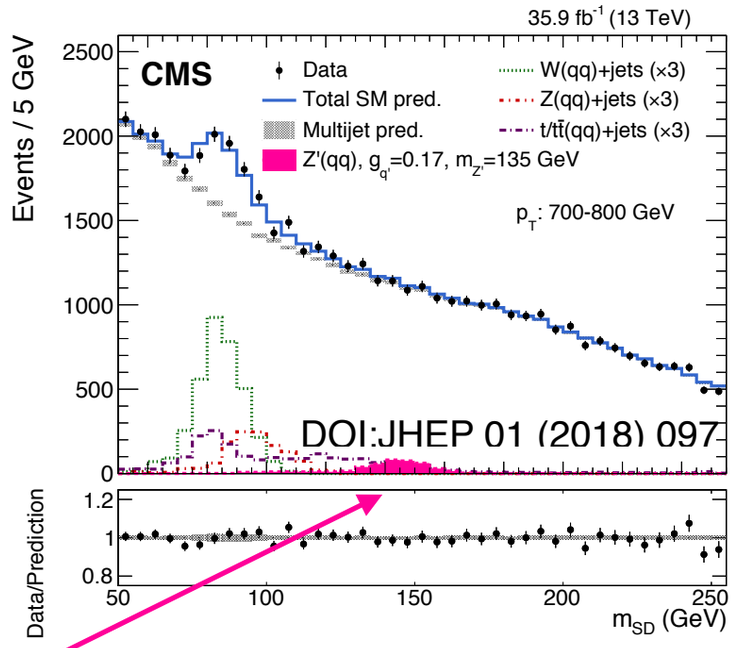
use full jet mass line shape, can incorporate all VV searches in one common framework

No signs of new physics in diboson searches

3D fit easily extendable to scan jet mass in search for non-SM bosons







Outlook: 2D → 3D

For all-hadronic, extend this to 3D scan of M_{V1}-M_{V2} - M_{VV}

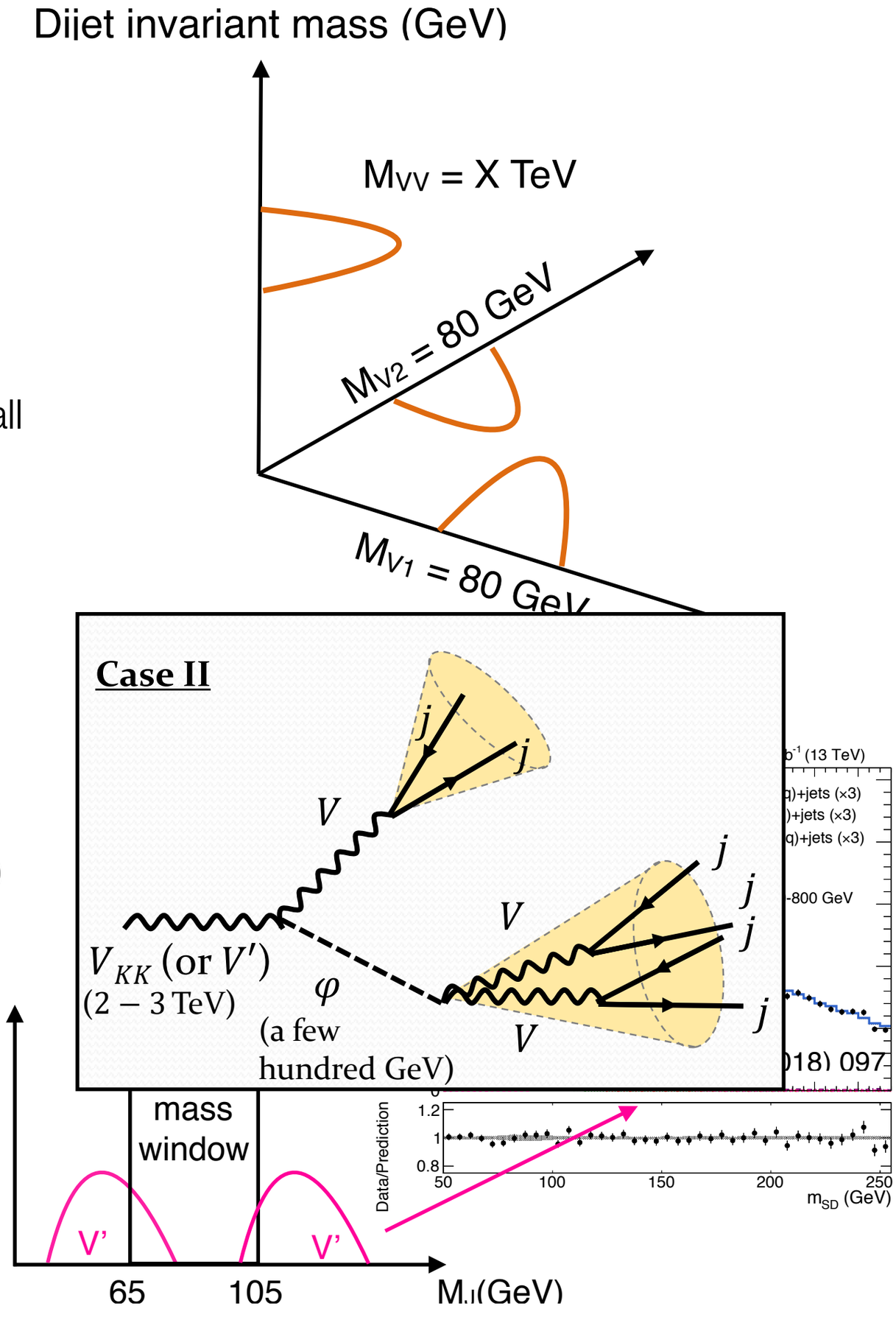
use full jet mass line shape, can incorporate all VV searches in one common framework

No signs of new physics in diboson searches

3D fit easily extendable to scan jet mass in search for non-SM bosons

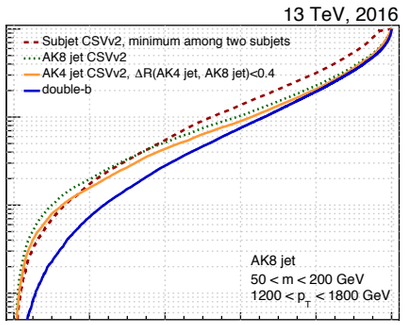
cascade decay signatures, scan different N-subjettines ratios (?), is softdrop optimal (?)

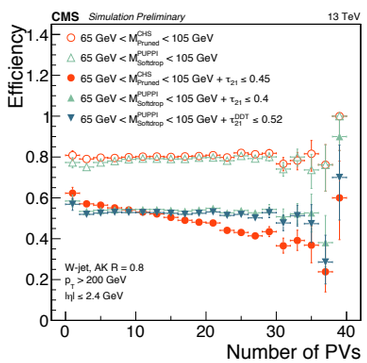
Ideal for searches with deep unsupervised learning and no signal priors (?) → learn from data!

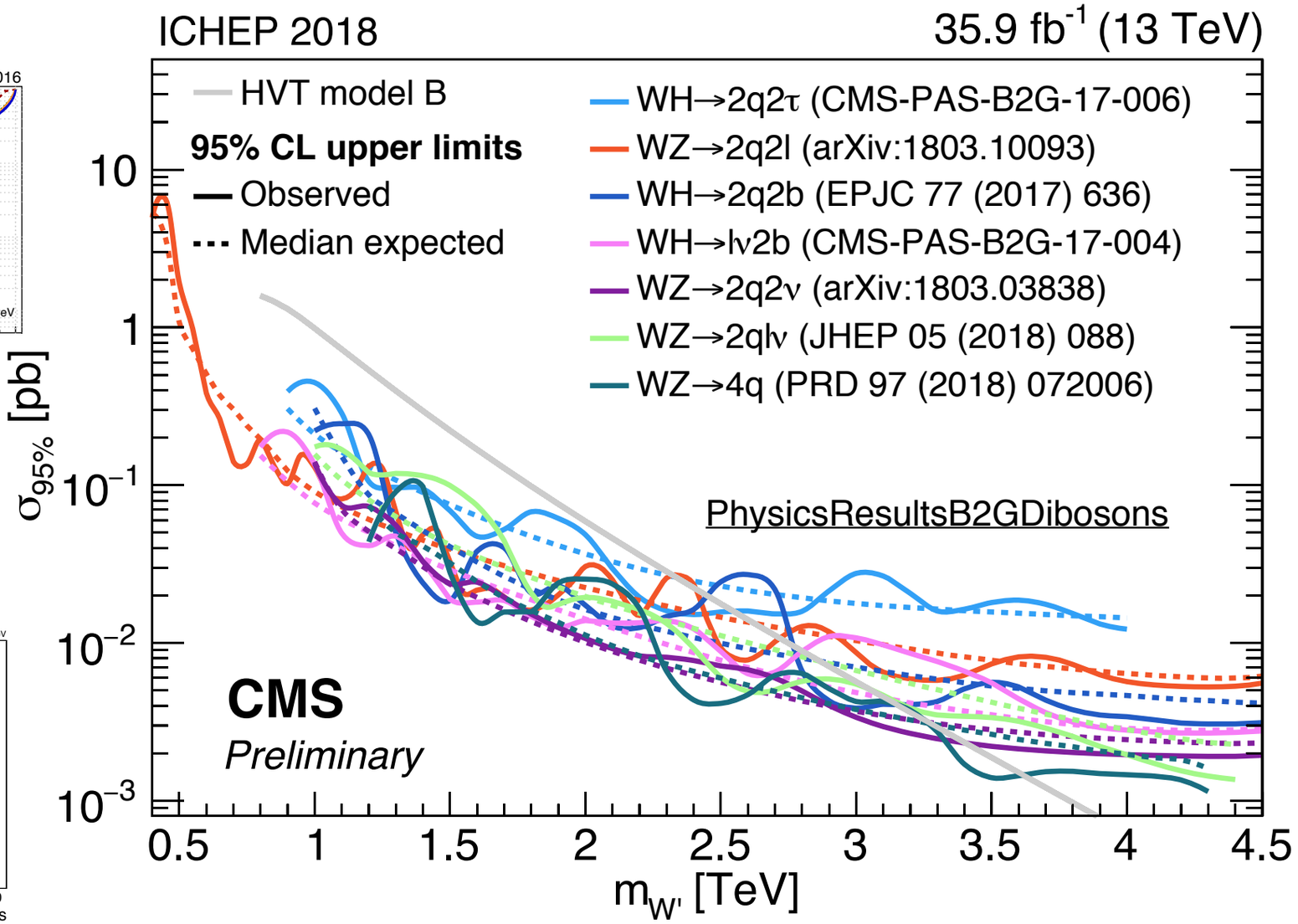


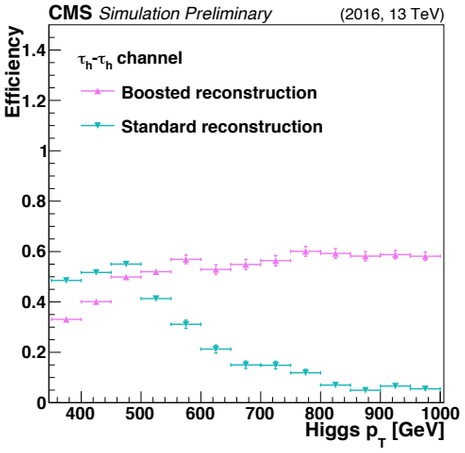
Summary

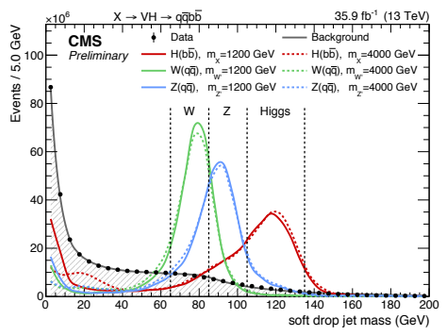
All lines on this plot represent dedicated methods boosted object reconstruction…

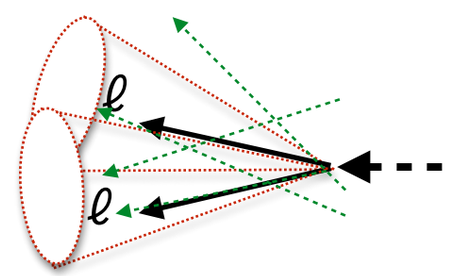








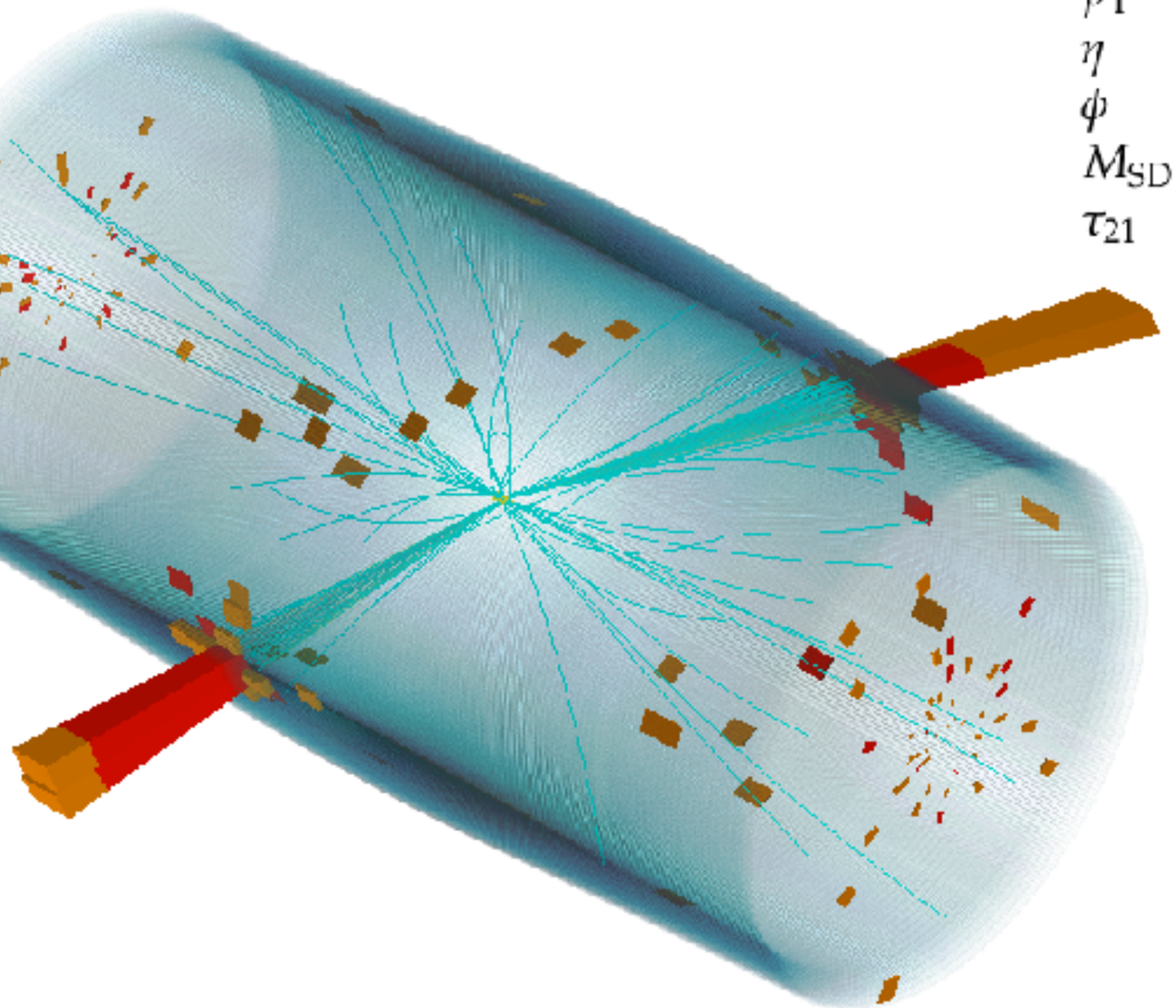




… novel tagging techniques and background estimation methods being commissioned to fully exploit the rich LHC Run 2 datasets!

Candidate ZZ event
Dijet mass: 3.2 TeV

Anti-k_T R=0.8 jet	
p_T	1374 GeV
η	0.79
ϕ	0.43
M_{SD}	94.8
τ_{21}	0.29



Anti-k_T R=0.8 jet	
p_T	1321 GeV
η	-0.40
ϕ	-2.71
M_{SD}	103 GeV
τ_{21}	0.23

CMS Experiment at LHC, CERN
Data recorded: Mon Jul 18 19:59:10 2016 CEST
Run/Event: 276950 / 1080730125
Lumi section: 573

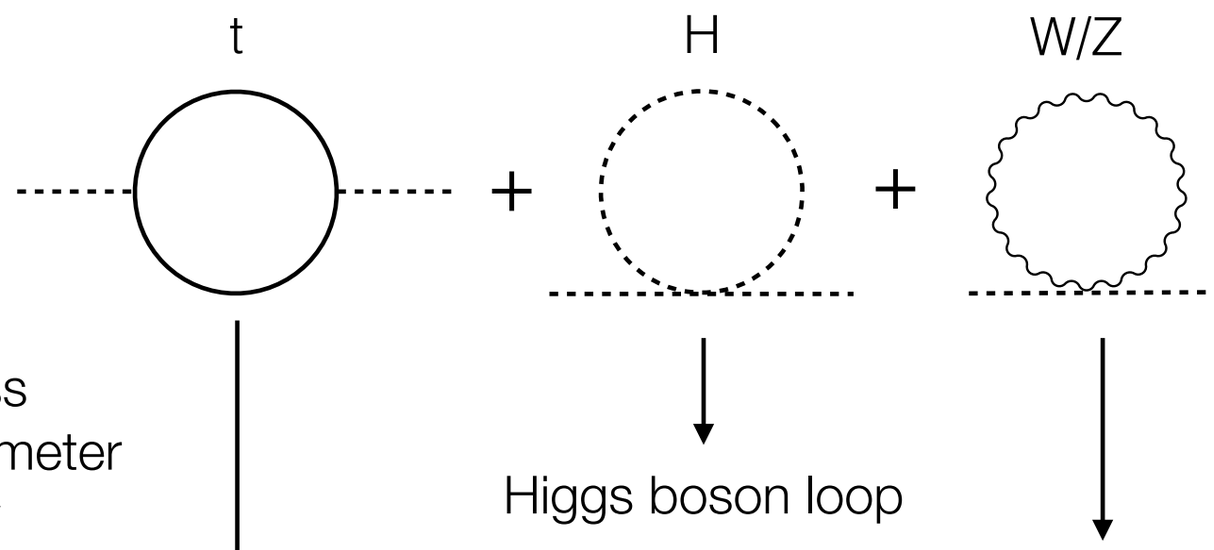
Backup slides

The hierarchy problem

quantum corrections $\propto M^2_{Planck} = 10^{38}$ GeV

physical value

$(125\ \text{GeV})^2 = m_H^2 -$



tree-level mass
fundamental parameter
of the theory

top quark loop

Higgs boson loop

gauge boson loop

Λ = energy scale up to which the SM remains valid = $O(M_{Planck})$

top loop $\quad -\frac{3}{8\pi^2}\lambda_t\Lambda^2$

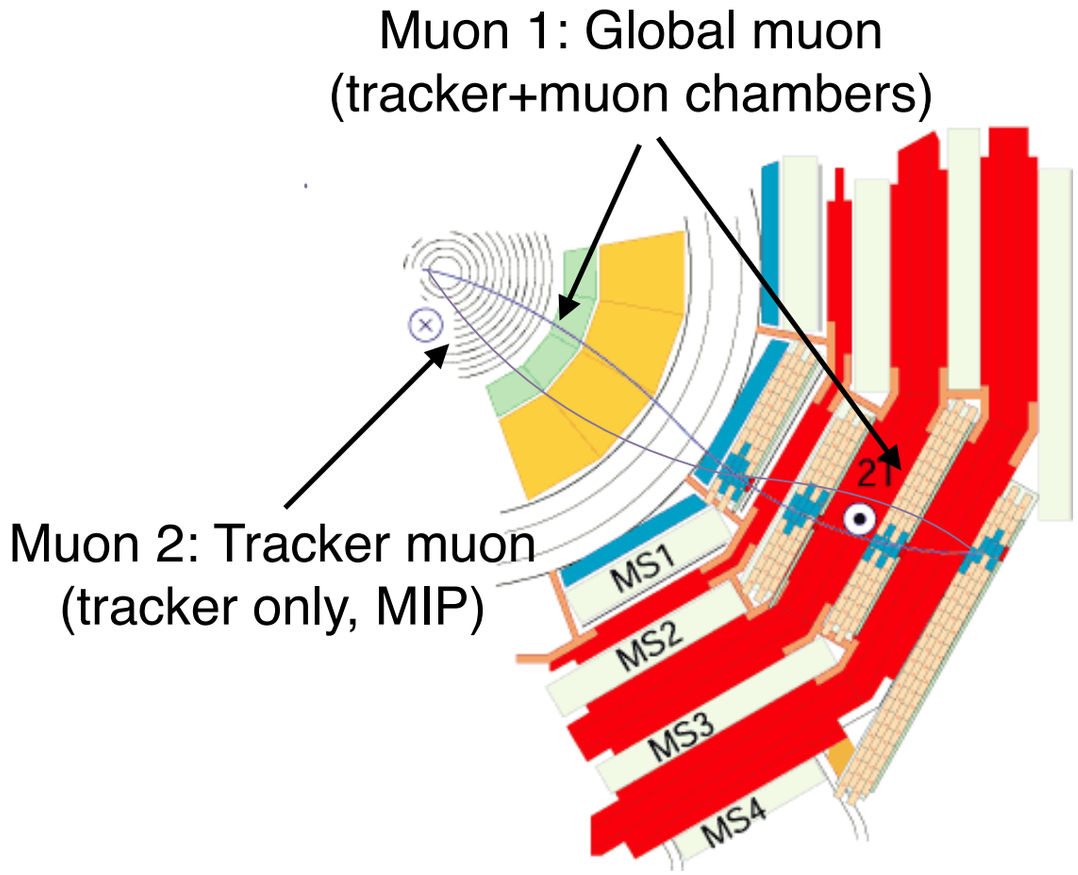
gauge loop $\quad +\frac{1}{16\pi^2}g^2\Lambda^2$

Higgs loop $\quad +\frac{1}{16\pi^2}\lambda^2\Lambda^2$

Boosted lepton reconstruction

For muons close in ΔR, “global muon” efficiency drops when same muon track segments are used to seed algorithm, only one muon left after cleaning

allow one muon to be identified in tracker only. Efficiency increase of 4-18%



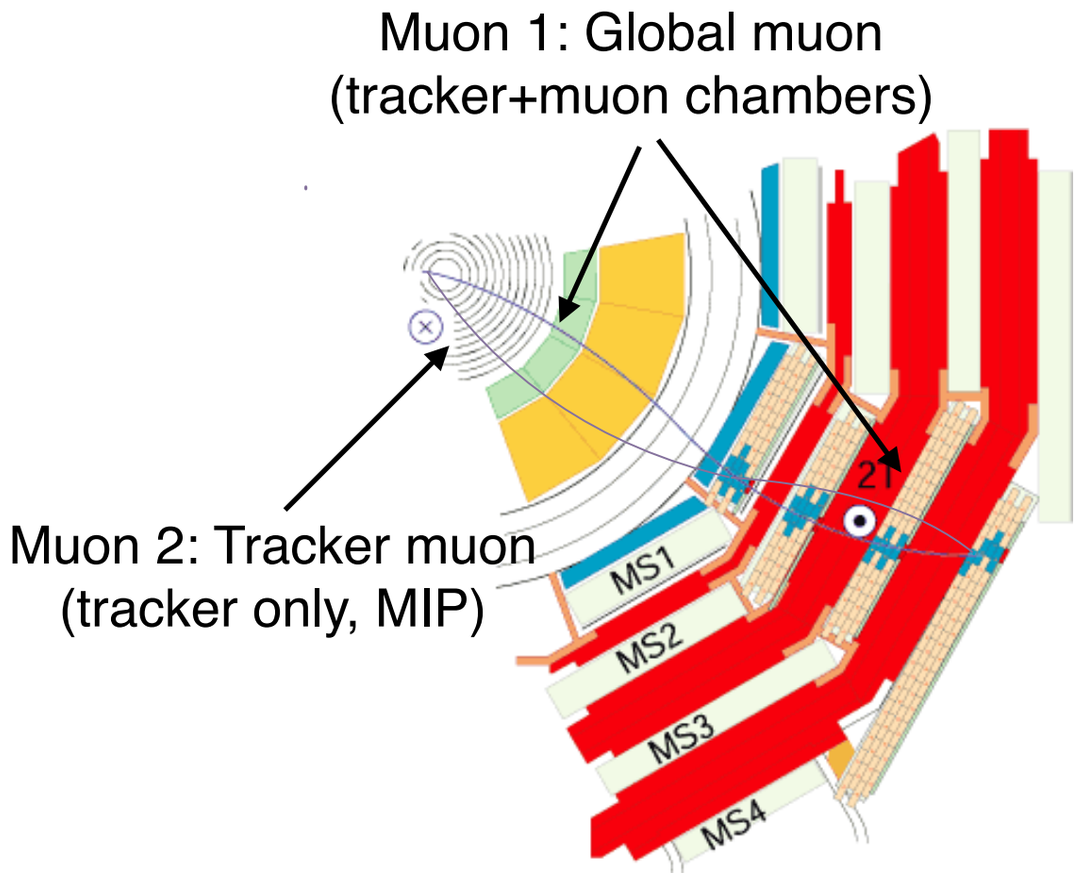
Boosted lepton reconstruction

For muons close in ΔR, “global muon” efficiency drops when same muon track segments are used to seed algorithm, only one muon left after cleaning

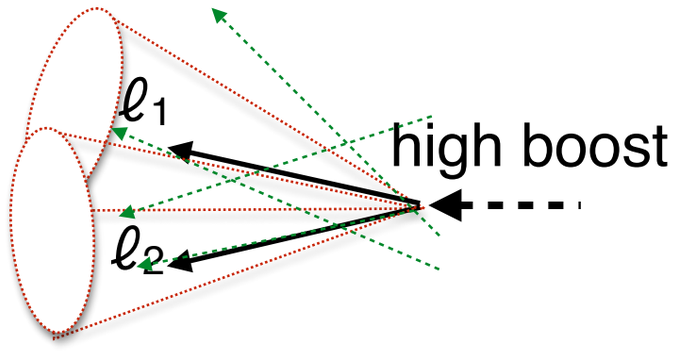
allow one muon to be identified in tracker only. Efficiency increase of 4-18%

Leptons required to be isolated from other activity in event

require lepton energy >> surrounding activity ($\sum p_T$ of charged particles within cone)



ℓ_1 iso: $\sum p_T / p_T\ell_1 < 10\%$



Requirement fails if second high-p_T lepton enters cone!

Boosted lepton reconstruction

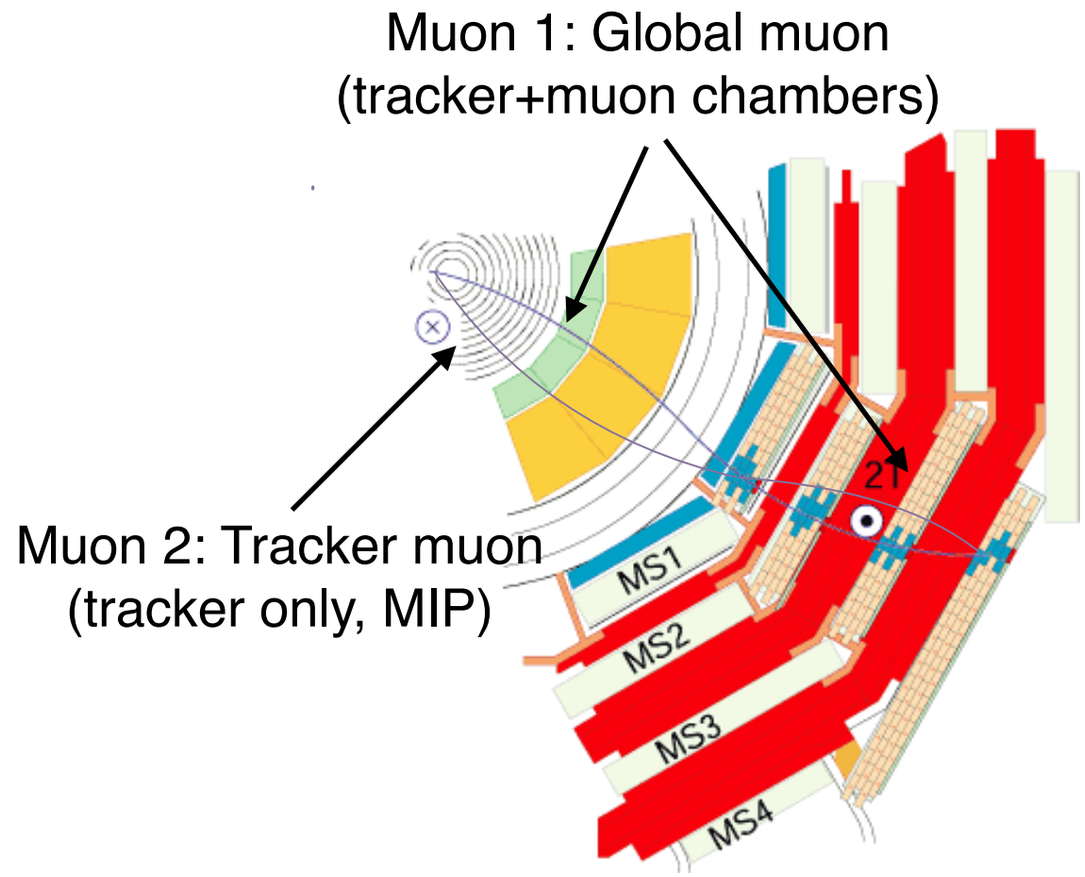
For muons close in ΔR, “global muon” efficiency drops when same muon track segments are used to seed algorithm, only one muon left after cleaning

allow one muon to be identified in tracker only. Efficiency increase of 4-18%

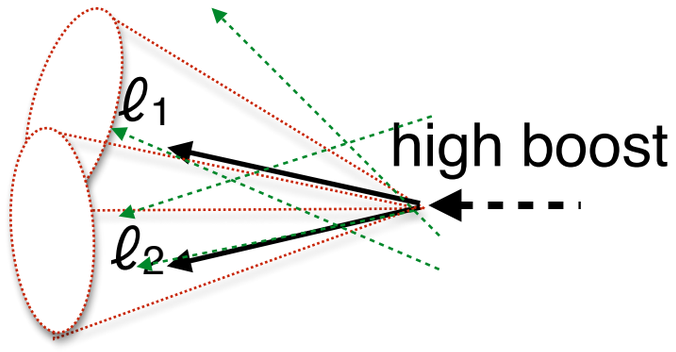
Leptons required to be isolated from other activity in event

require lepton energy >> surrounding activity ($\sum p_T$ of charged particles within cone)

to retain efficiency at high boost, remove all other PF electrons and muons before computing isolation



ℓ_1 iso: ($\sum p_T$ - $p_{T,\ell 2}$) / $p_{T\ell 1}$ < 10%



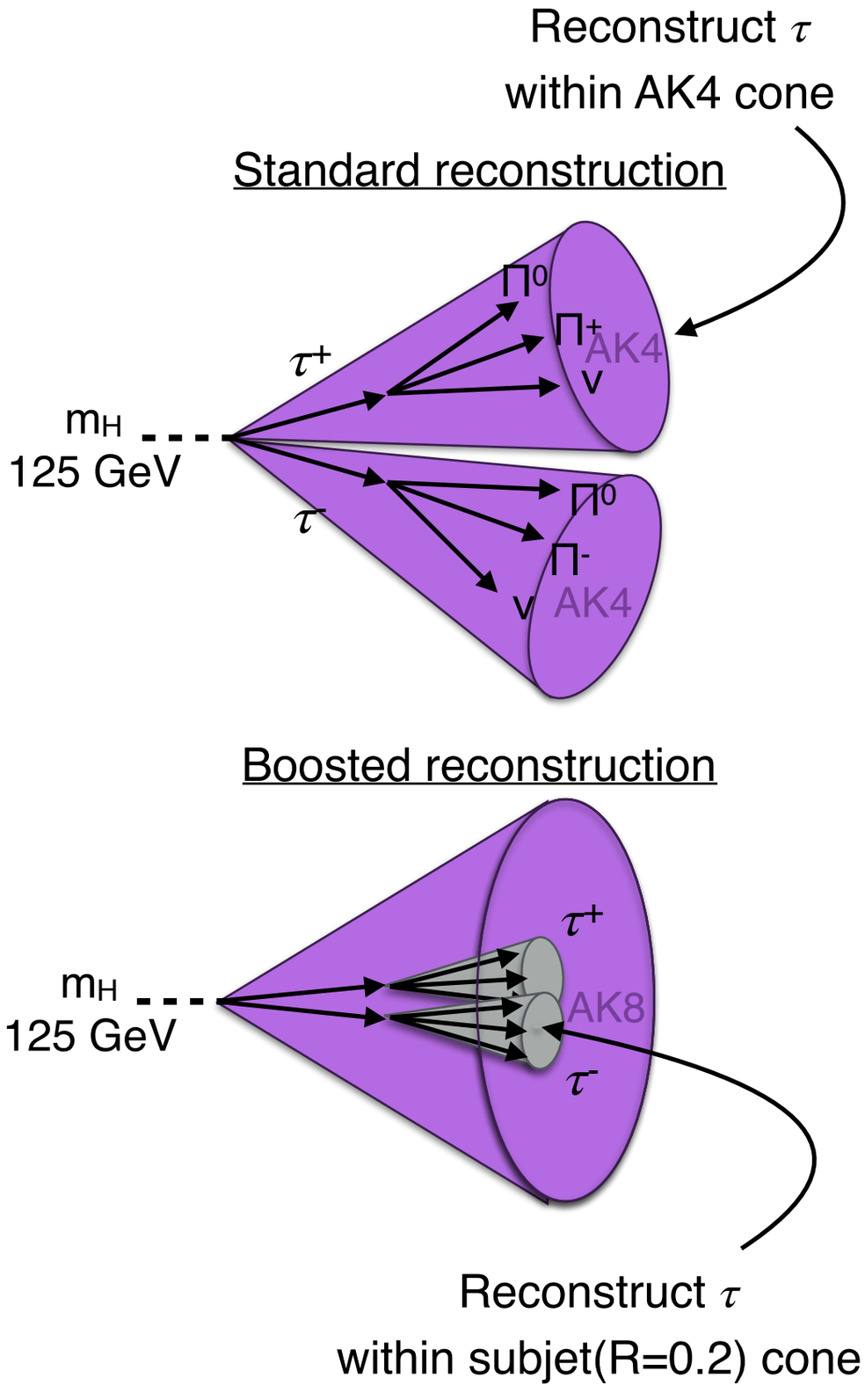
Boosted τ reconstruction

Standard τ reconstruction algorithms attempt to identify τ by reconstructing τ decay products (π^0,$\pi^{+/-}$)

look within AK4 jet cone

For boosted $\tau\tau$ final states, AK4 cones overlap and efficiency drops

Rather "seed" reconstruction using AK8 softdrop subjets R=0.2 → subjet τ-tagging



Boosted τ reconstruction

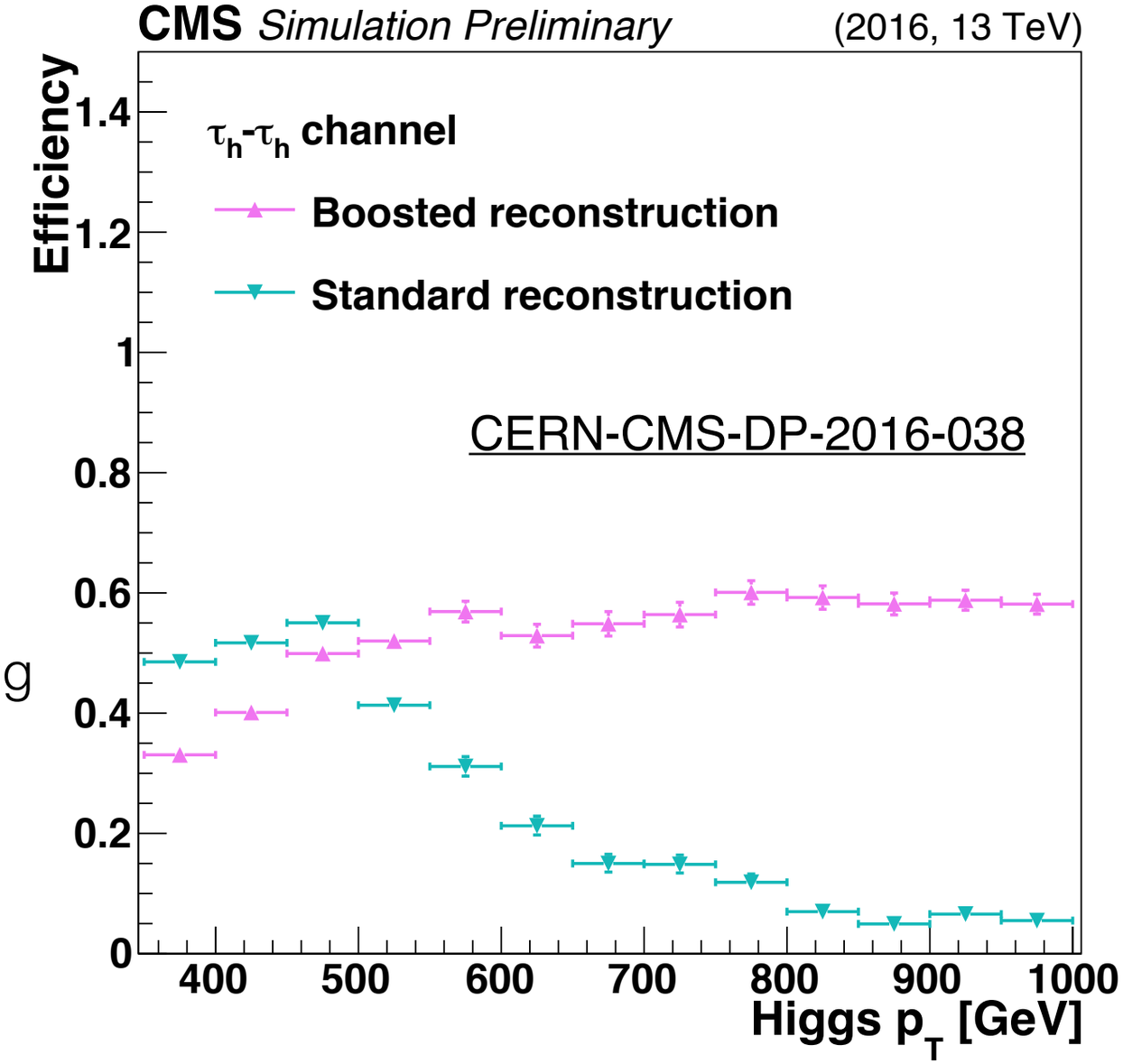
Standard τ reconstruction algorithms attempt to identify τ by reconstructing τ decay products (π^0,$\pi^{+/-}$)

look within AK4 jet cone

For boosted $\tau\tau$ final states, AK4 cones overlap and efficiency drops

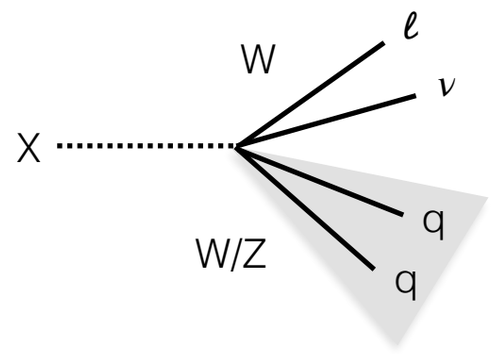
Rather “seed” reconstruction using AK8 softdrop subjets R=0.2 → subjet τ-tagging

Huge gain in efficiency using dedicated reconstruction!



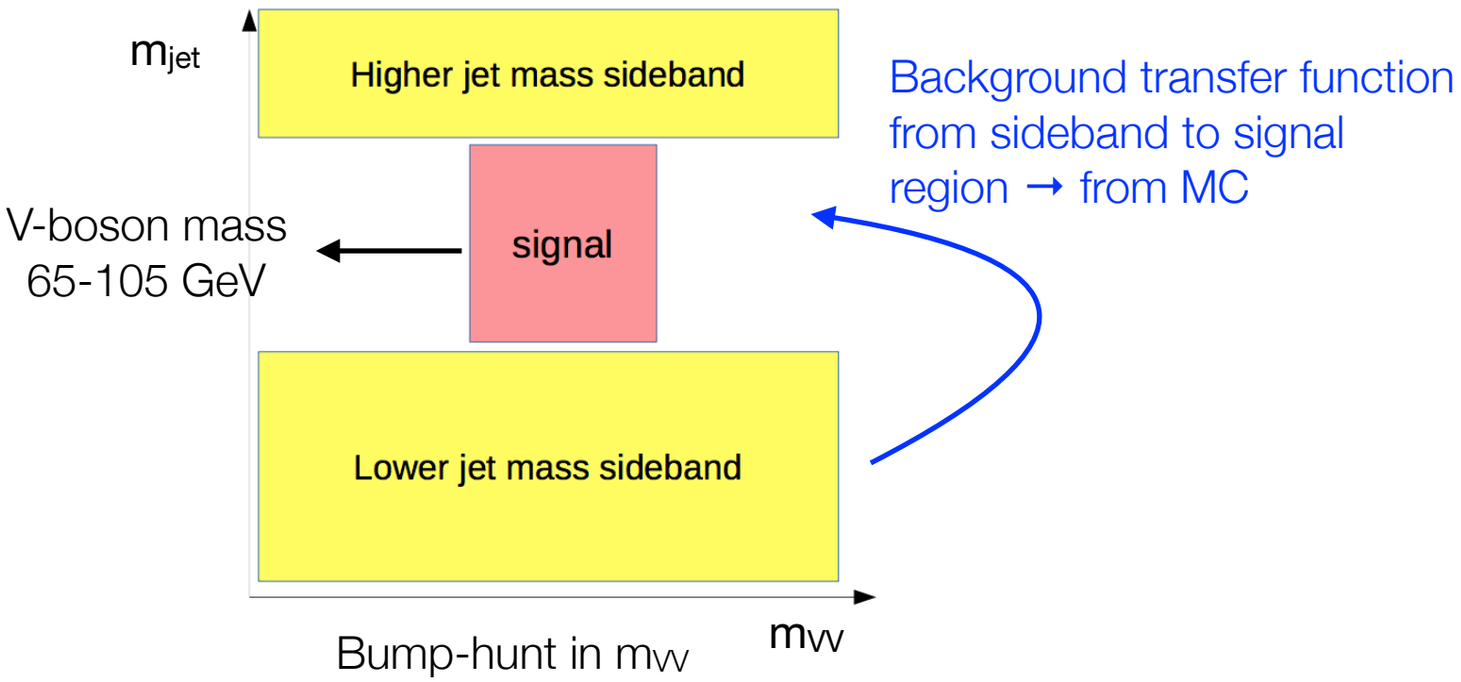
2D fit: X→WV→ℓνqq

- Event topology: 1 isolated lepton + E_T^{miss} + high p_T AK8 jet (V-jet)
- Discriminating variables: jet softdrop mass (m_{jet}) and $m_{\ell\nu+jet}$ (m_{VV})



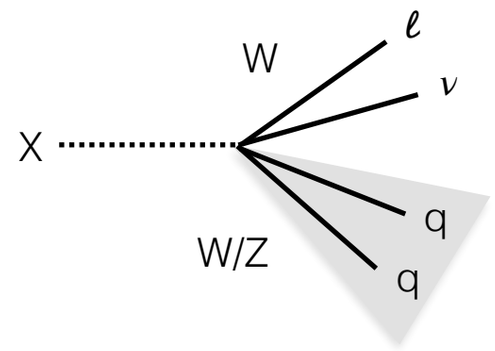
Old 1D method (α-method):
Mass-sideband aided bump hunt

- estimate m_{VV} in the signal region from data in m_{jet} sideband
- default and well validated method since Run 1



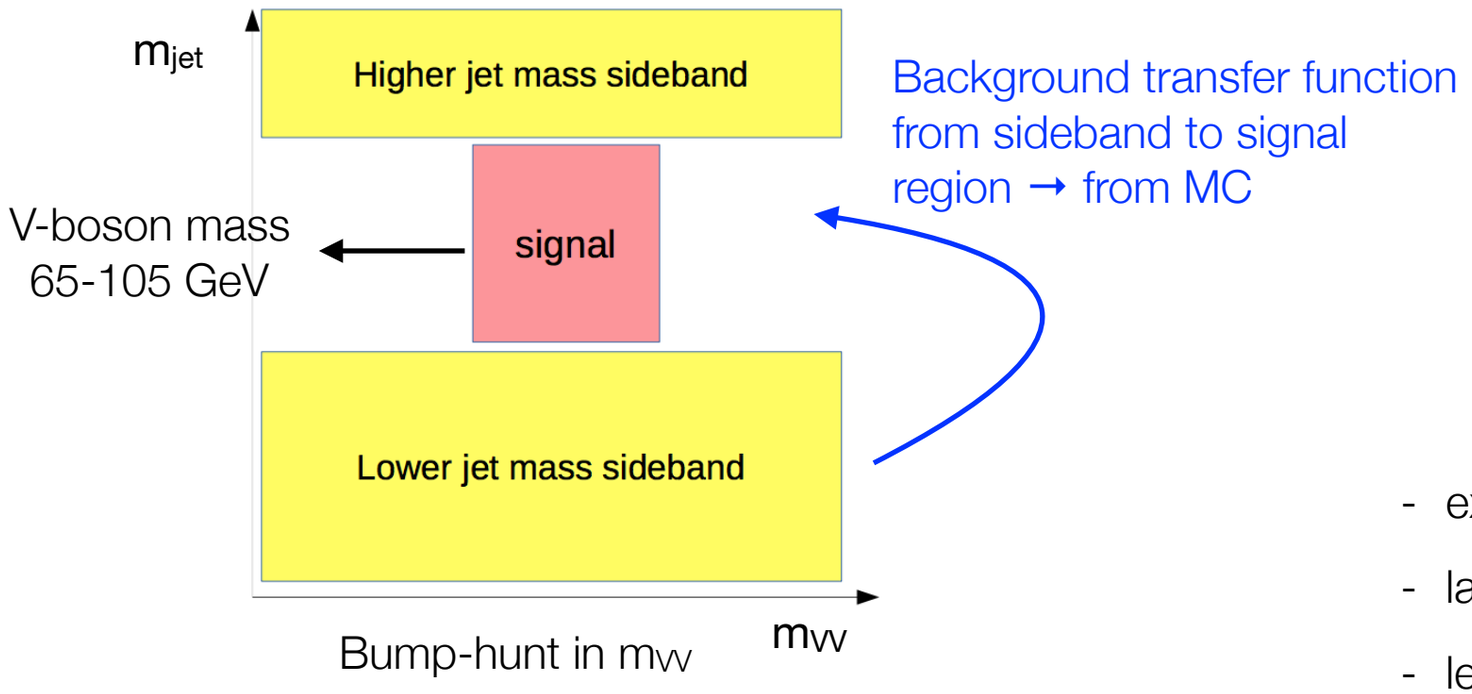
2D fit: X→WV→ℓνqq

- Event topology: 1 isolated lepton + E_T^{miss} + high p_T AK8 jet (V-jet)
- Discriminating variables: jet softdrop mass (m_{jet}) and $m_{\ell\nu+jet}$ (m_{VV})

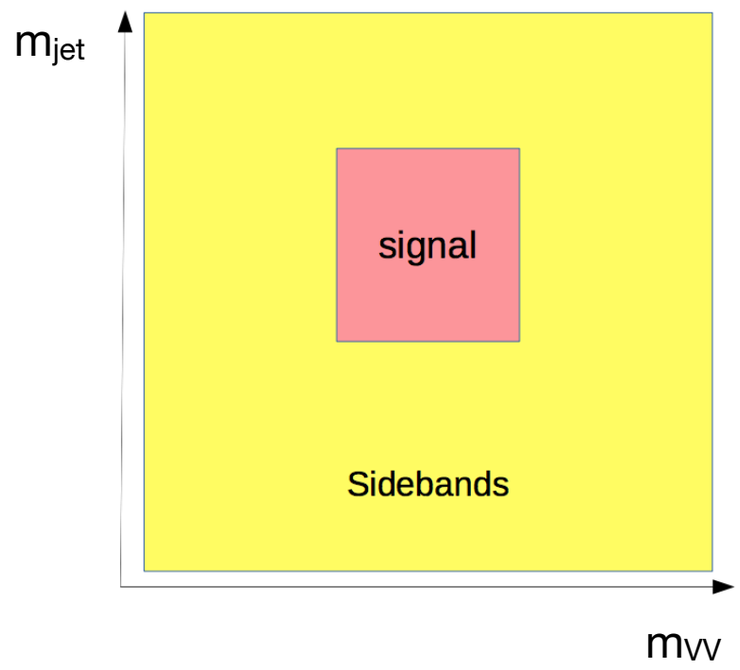


Old 1D method (α-method):
Mass-sideband aided bump hunt

- estimate m_{VV} in the signal region from data in m_{jet} sideband
- default and well validated method since Run 1



NEW 2D method:
2D bump hunt in the (m_{jet},m_{VV}) plane
using full mjet range



B2G-16-029

- exploit correlations between m_{VV} and m_{jet}
- larger statistics and use full line-shape of jet mass
- less dependent on simulation and learn from data

1D background estimation

⇒ estimate this background from data in signal-free sidebands around the m_{jet} signal window

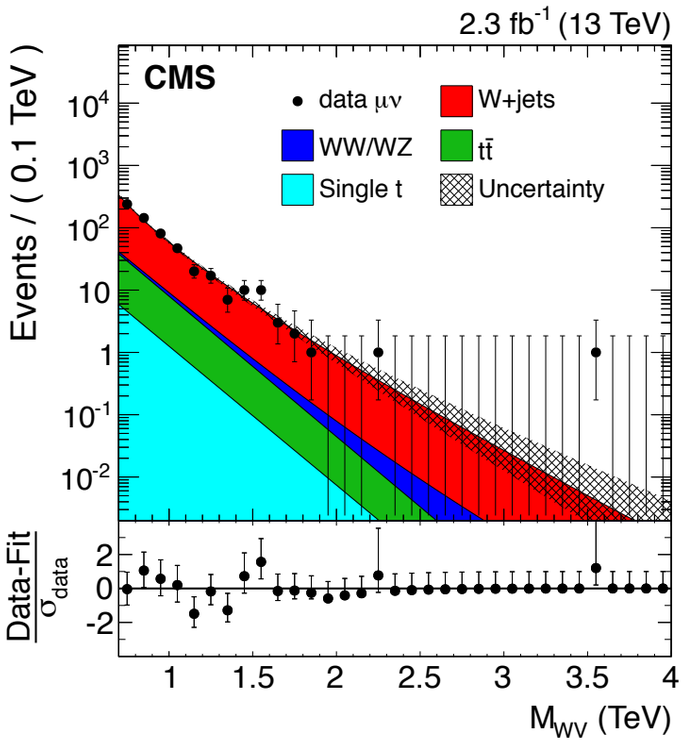
1) fit data in sideband

with exponentially falling function

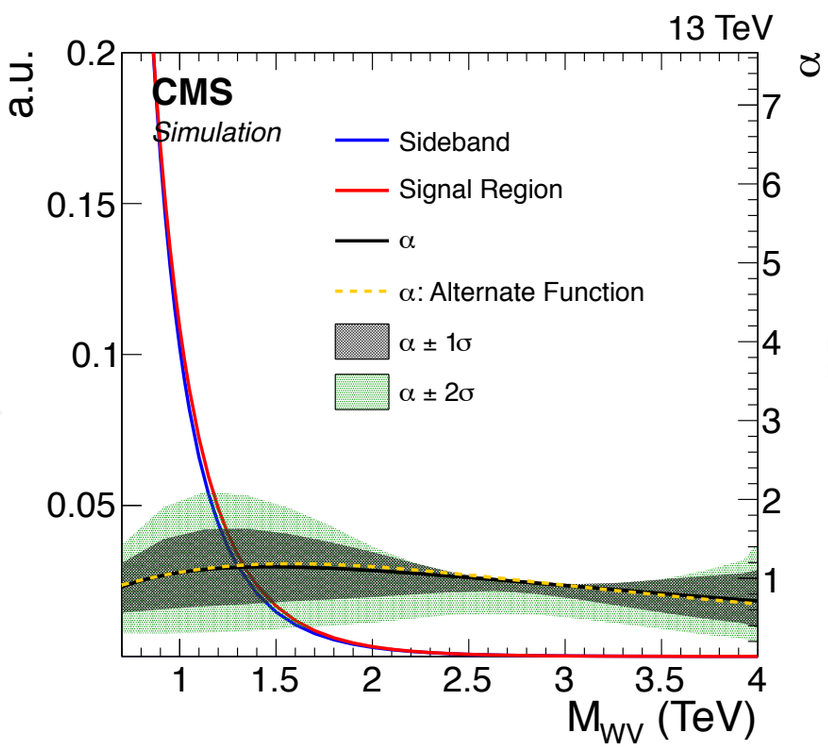
2) compute transfer function from simulation

$$\alpha_{MC} = \frac{F_{MC}\ (\text{signal region})}{F_{MC}\ (\text{sideband})}$$

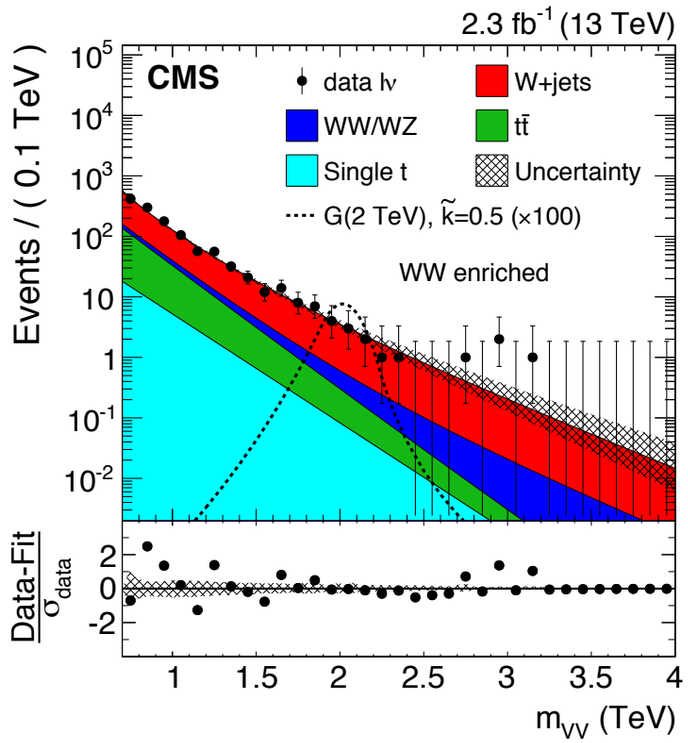
3) predict background in the signal region



X



=



Top quark and non-resonant diboson backgrounds from simulation (corrected)