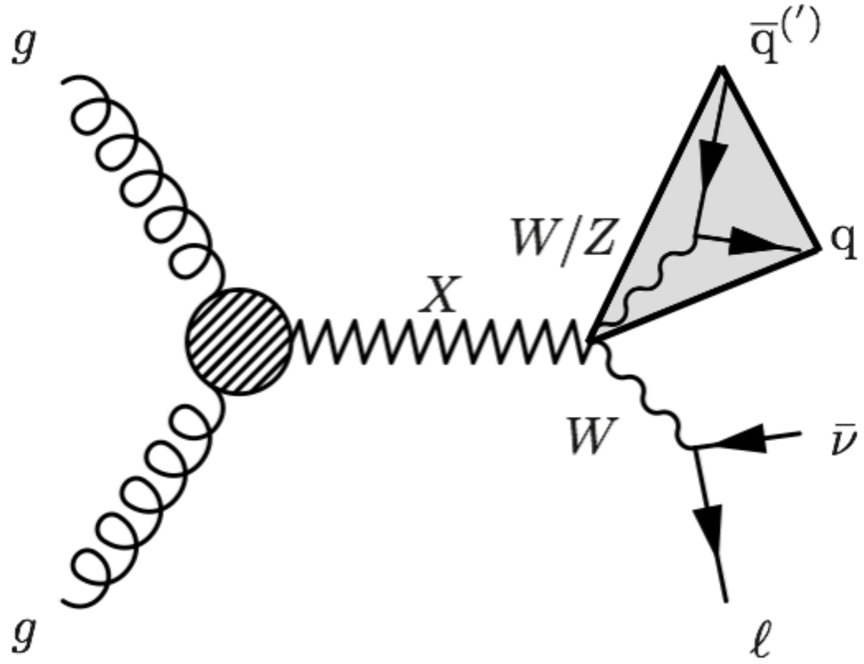
# Search for heavy resonances in diboson final states at CMS

## Highlights of the CMS diboson resonances search programme



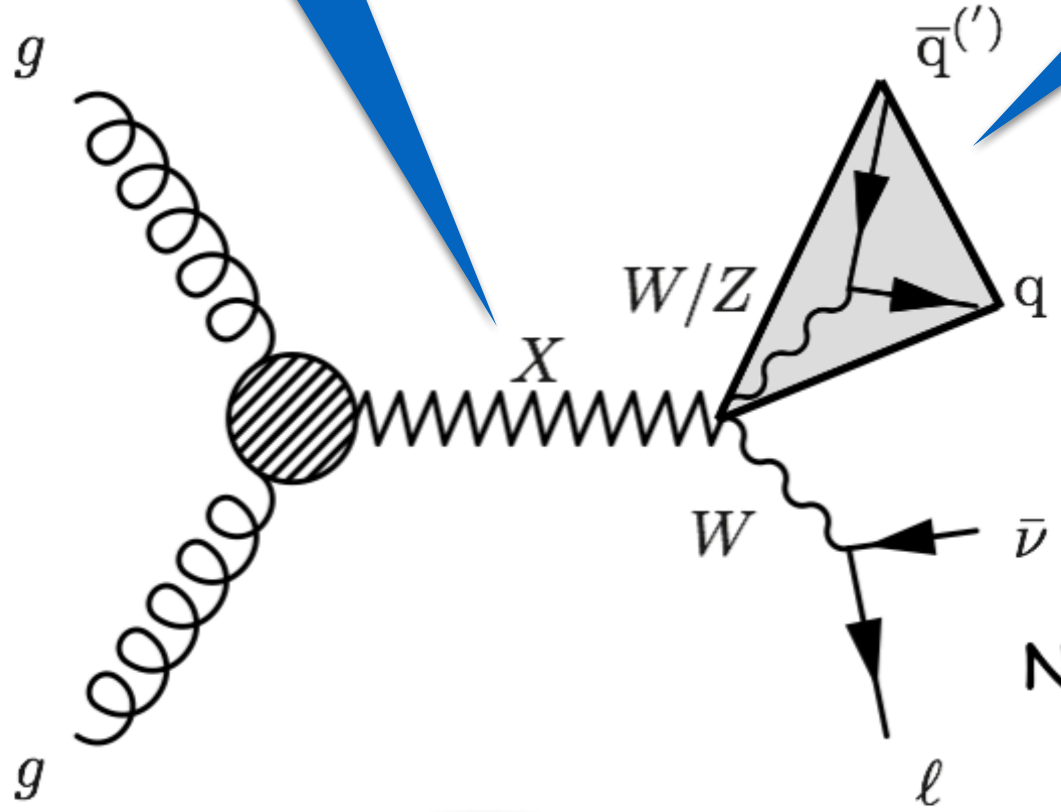
Clemens Lange (CERN)
on behalf of the CMS Collaboration

CERN-LHC Seminar
5th December 2017

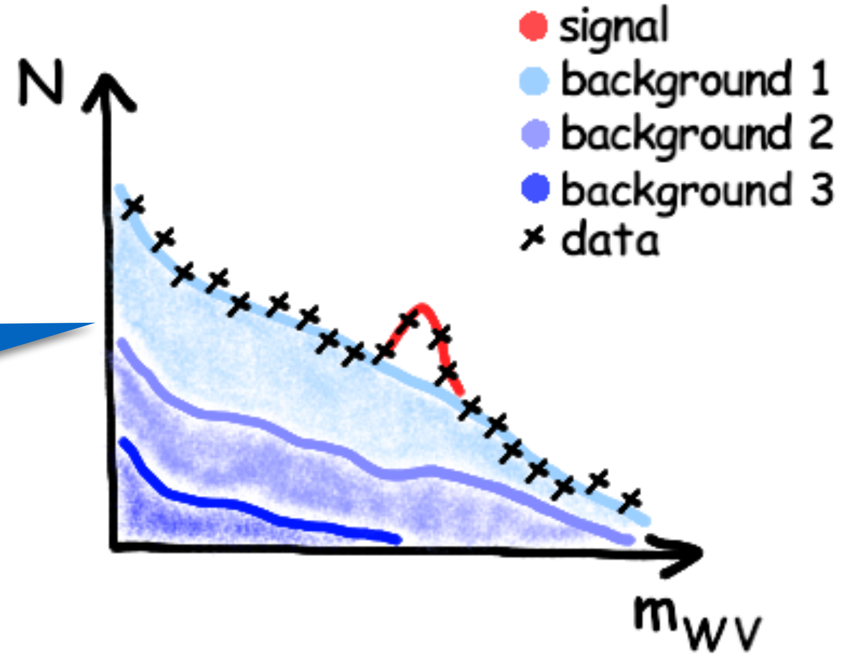
# Diboson resonances

heavy resonances?

reconstruction of bosons?

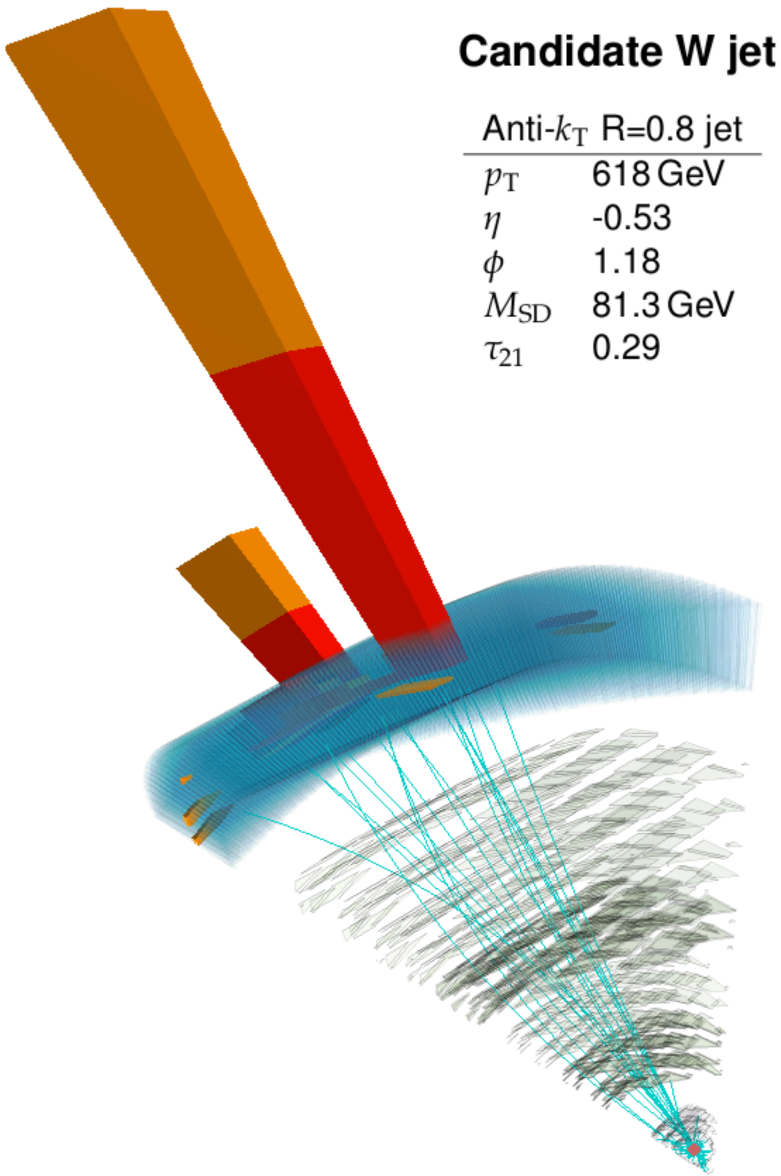


background estimation?



# Outline

- Why search for diboson resonances?
- Boson reconstruction
  - jet substructure
- Diboson resonance searches
  - all-hadronic VV/Vh final states
  - di-Higgs (hh) final states
  - (semi-)leptonic VV final states



**Candidate W jet**

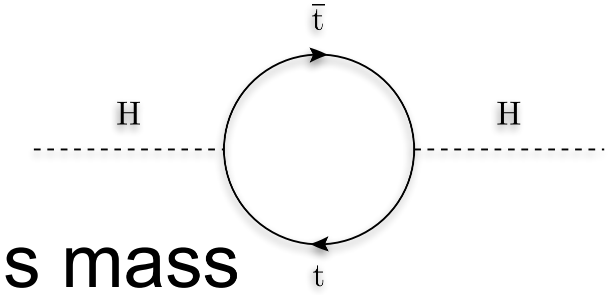
| Anti-$k_T$ R=0.8 jet | |
|---|---|
| $p_T$ | 618 GeV |
| $\eta$ | -0.53 |
| $\phi$ | 1.18 |
| $M_{SD}$ | 81.3 GeV |
| $\tau_{21}$ | 0.29 |

**Results of 4 analyses shown publicly for the first time today**

# The hierarchy problem

- how can we explain the big difference between the **weak force** and **gravitation**?

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

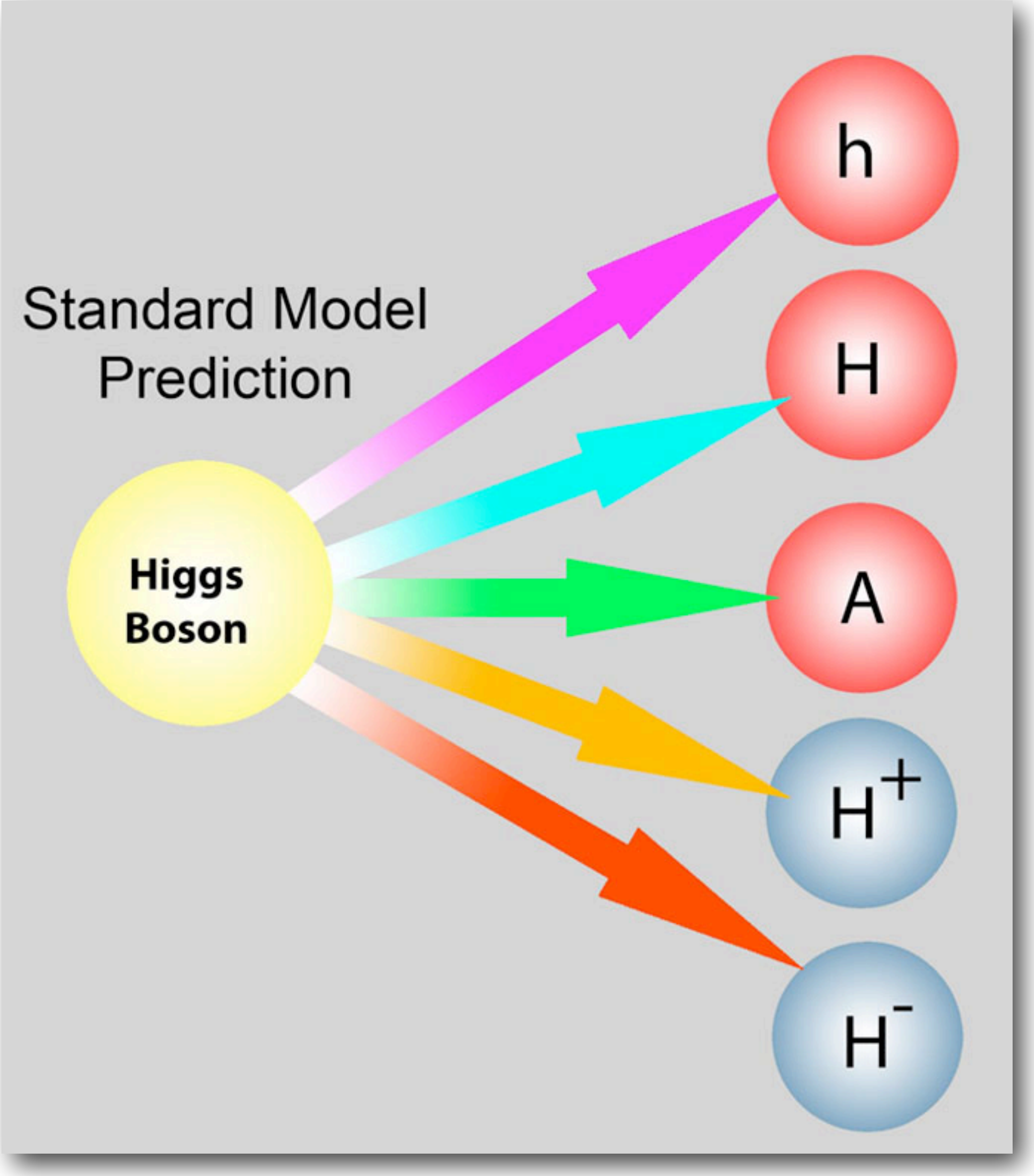


- no **symmetry** in the standard model (SM) protects the Higgs mass
- „natural“ explanation would be that SM is replaced/extended by another theory at the TeV scale: $\mu^2 \sim (\text{heavier scale})^2$ ➜ **new particles**

- these theories could be (among others):
- **SUSY**: protecting the Higgs mass by a symmetry
- **Composite Higgs**: the Higgs is not elementary
- **Large/warped extra dimensions**: gravity is strong at electroweak scale

# 2-Higgs-Doublet-Models

- 2-Higgs-Doublet-Models (2HDM) one of the simplest extensions of the SM
  - motivation from SUSY, axion models, baryon asymmetry, ...
- usually two complex scalar SU(2) doublets leading to five scalar Higgs fields:
  - charged scalar ($H^\pm$), two neutral scalars (H and h), one pseudoscalar (A)
- Here: light scalar (h) 125 GeV boson
  - heavier particles' mass > $2\times m_W/m_Z/m_H$

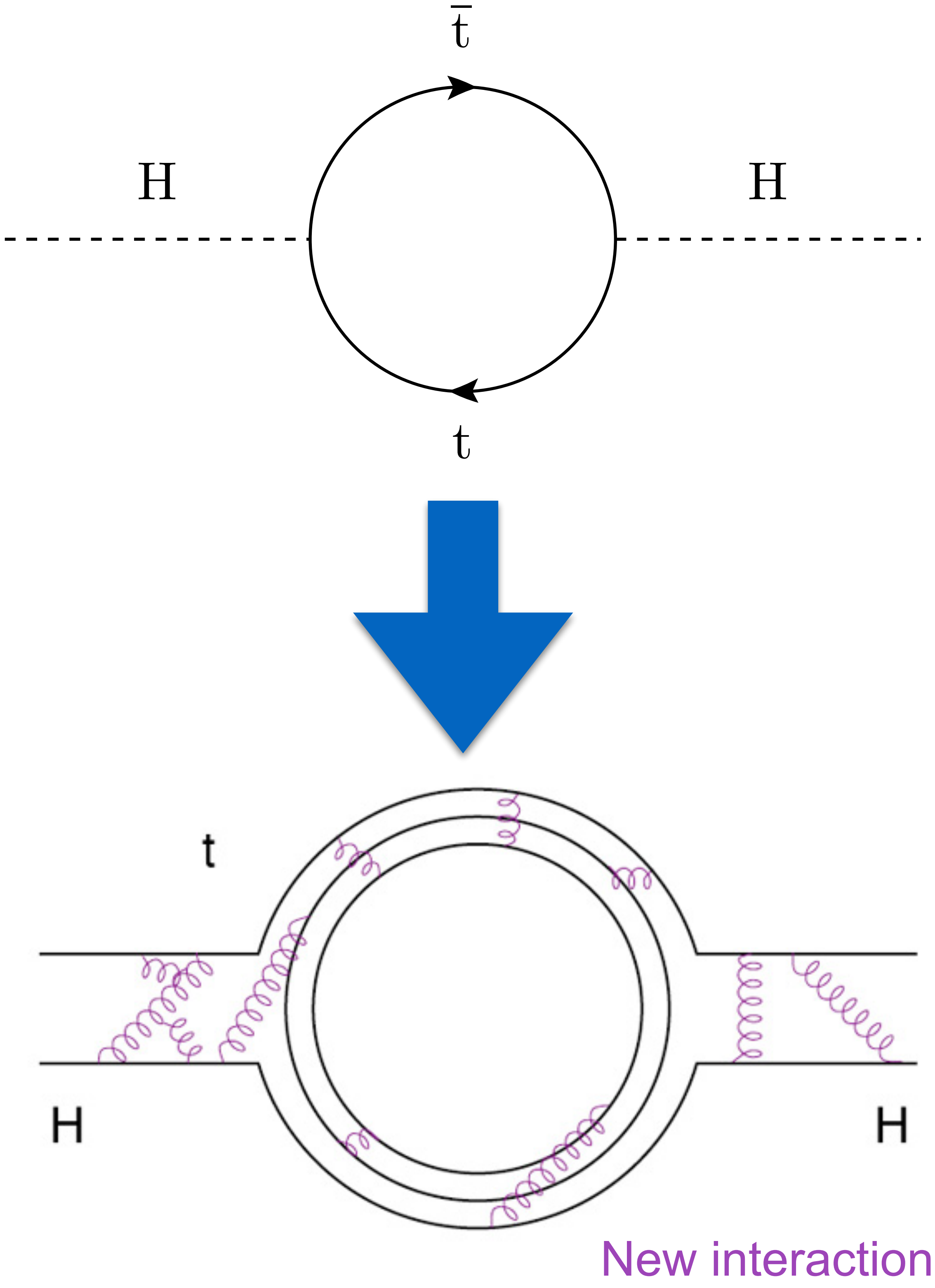


**Here:**
**mostly scalar (spin-0) resonances decaying to pairs of SM h bosons (decays into W/Z pairs also possible)**

# Composite Higgs

www.learner.org

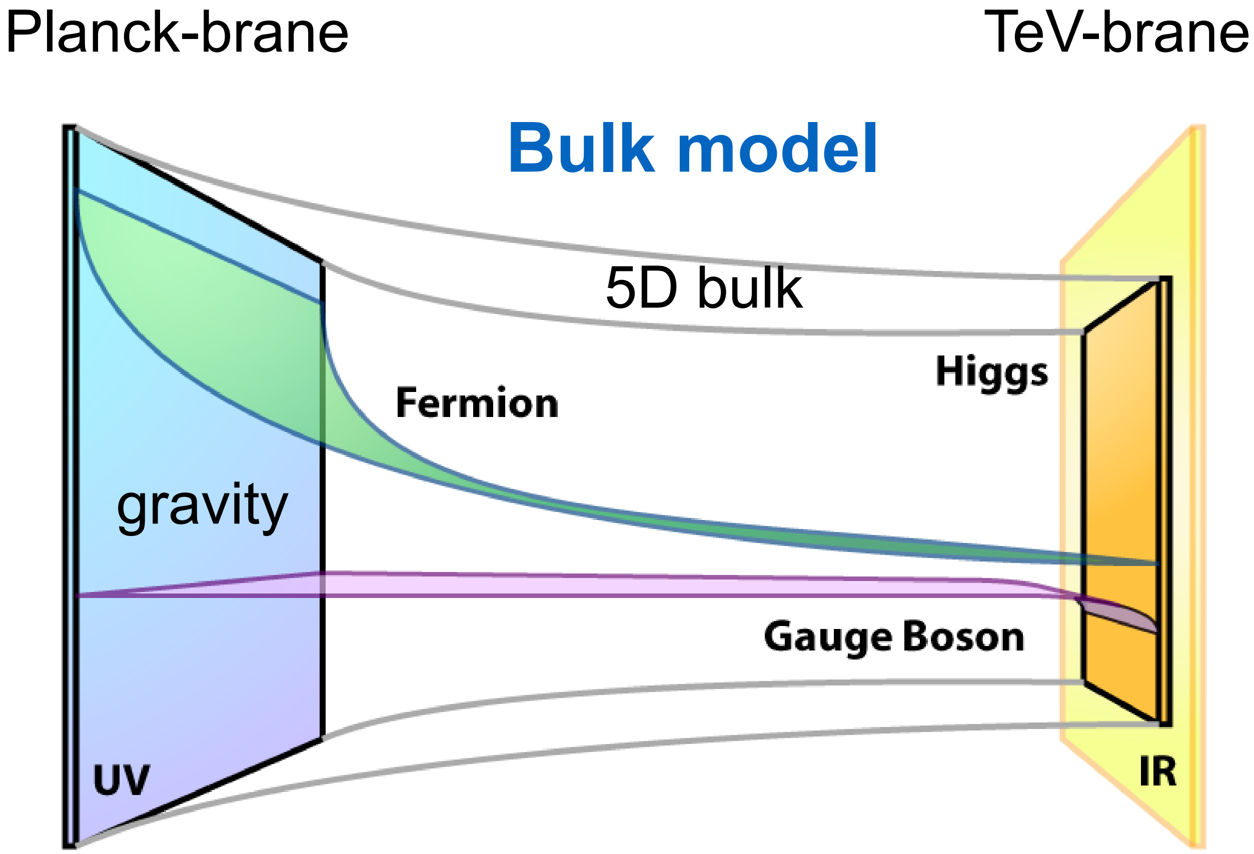
- The **Higgs** could be **non-fundamental**
- Instead: **bound state** of a new strong interaction
- Brings along **new heavy particles**/ states
- Heavy partners of SM particles decay to lighter ones (W, Z, h, top, …)



**Here:**
**spin-1 (W’, Z’) resonances**
**decaying to pairs/combinations of W, Z, h**
**interpretation in “heavy vector triplet” framework (W’/Z’, V’)**

# (Warped) extra dimensions?

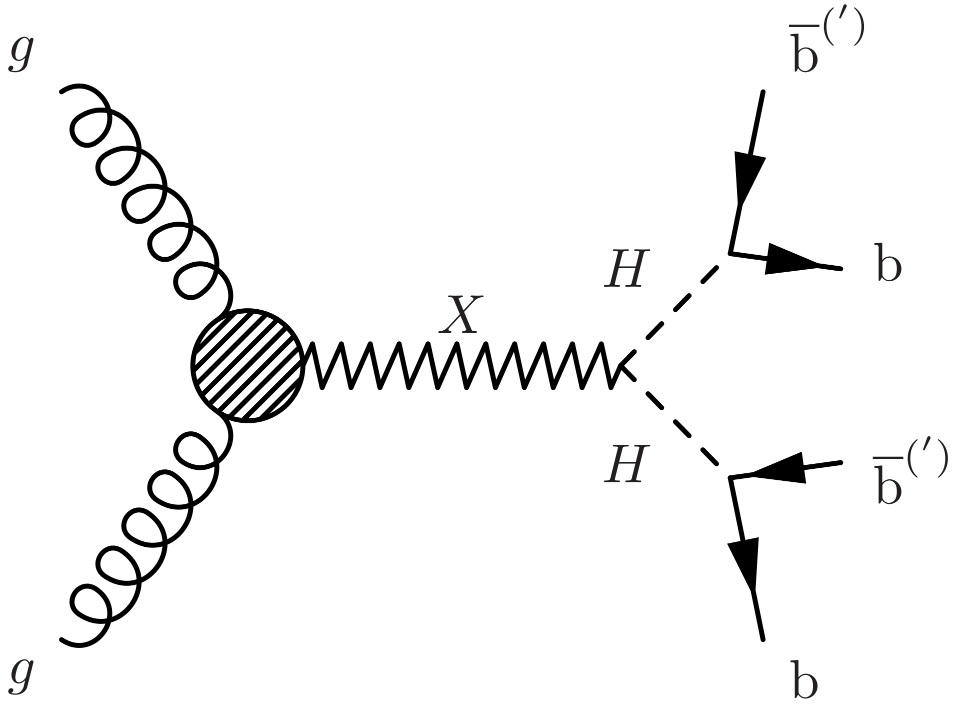
- SM fields are confined to **four-dimensional „membrane“**, gravity propagates in additional dimensions
- Change effective Planck constant, **reducing Planck scale** to close **to electroweak scale**
- Overlap of 5D profiles at TeV-brane (and Higgs) determine **particle masses**
- Additionally, if distance between two branes is not fixed, **additional fluctuations** can occur

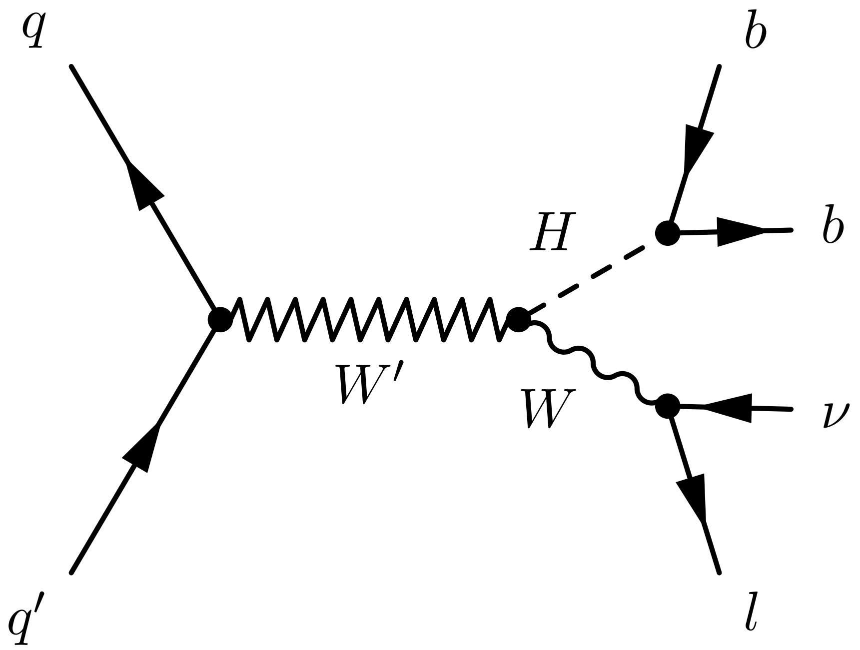


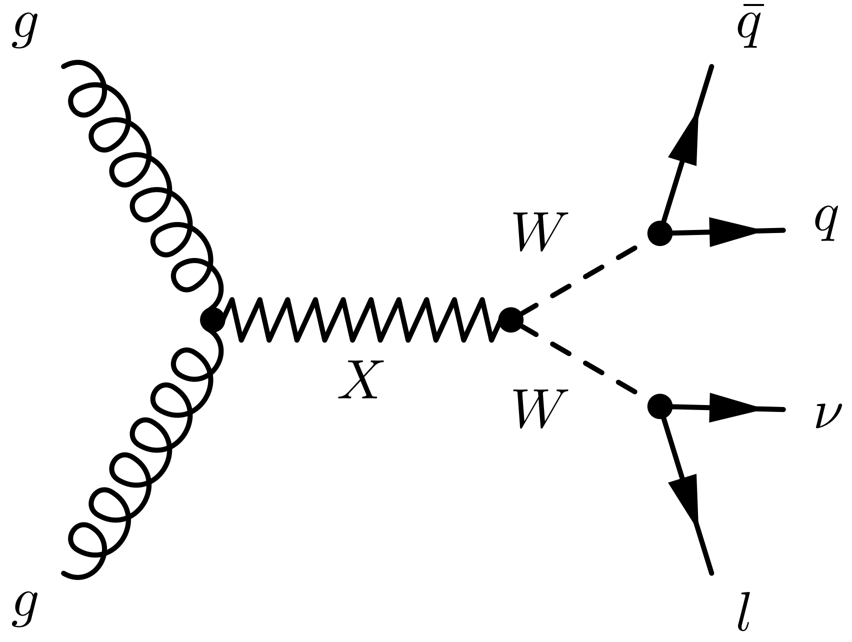
**Here:**
**spin-2 gravitons**
**decaying to pairs of W, Z, (ɣ)**
**—**
**spin-0 radion fluctuations**
**decaying to pairs of SM h bosons**

# How do we observe these signals at the LHC?

- In **diboson final states**, should be able to observe excitations/ resonances/fluctuations
- Spin 0, 1, or 2
- Depending on **model parameters**, resonances can be narrow or wide
  - model-independent interpretation important
- Majority of analyses presented here focus on **narrow resonances** (width < detector resolution), **mass ≥ 130 GeV**



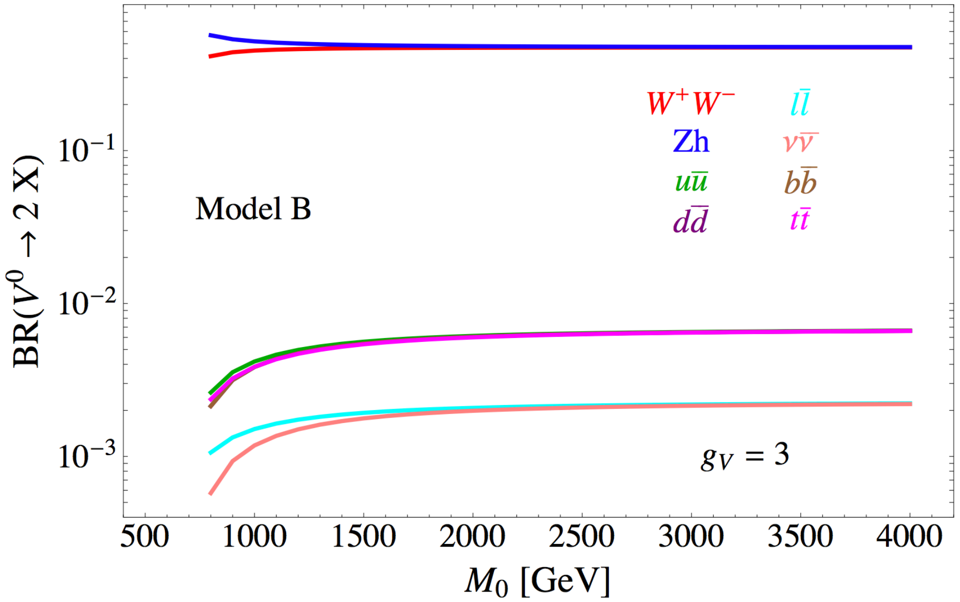


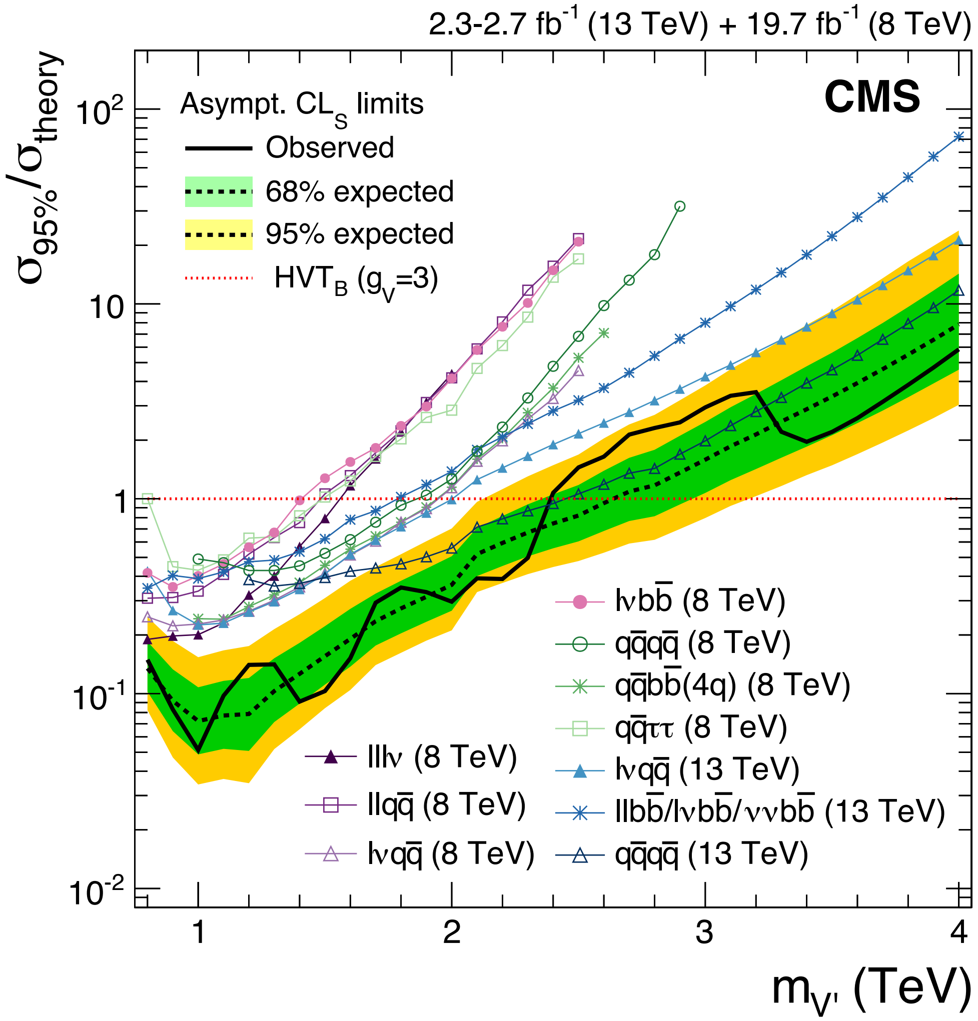


# Status last year

JHEP 1409 (2014) 060
B2G-16-007 (accepted by Phys. Lett. B)

- Performed **statistical combination** of all CMS √s = 8 TeV (2012) and 2015 13 TeV **diboson high mass analyses**
- Benchmark models:
  - Heavy vector triplet (HVT) models A and B
  - Bulk graviton model
- While individual analyses showed deviations from SM background expectation > 2σ, combination showed none
- Full **2016 data (35.9 fb$^{-1}$) analyses have significantly higher sensitivity**
  - signal cross sections 13 TeV to 8 TeV a factor 2–5 higher



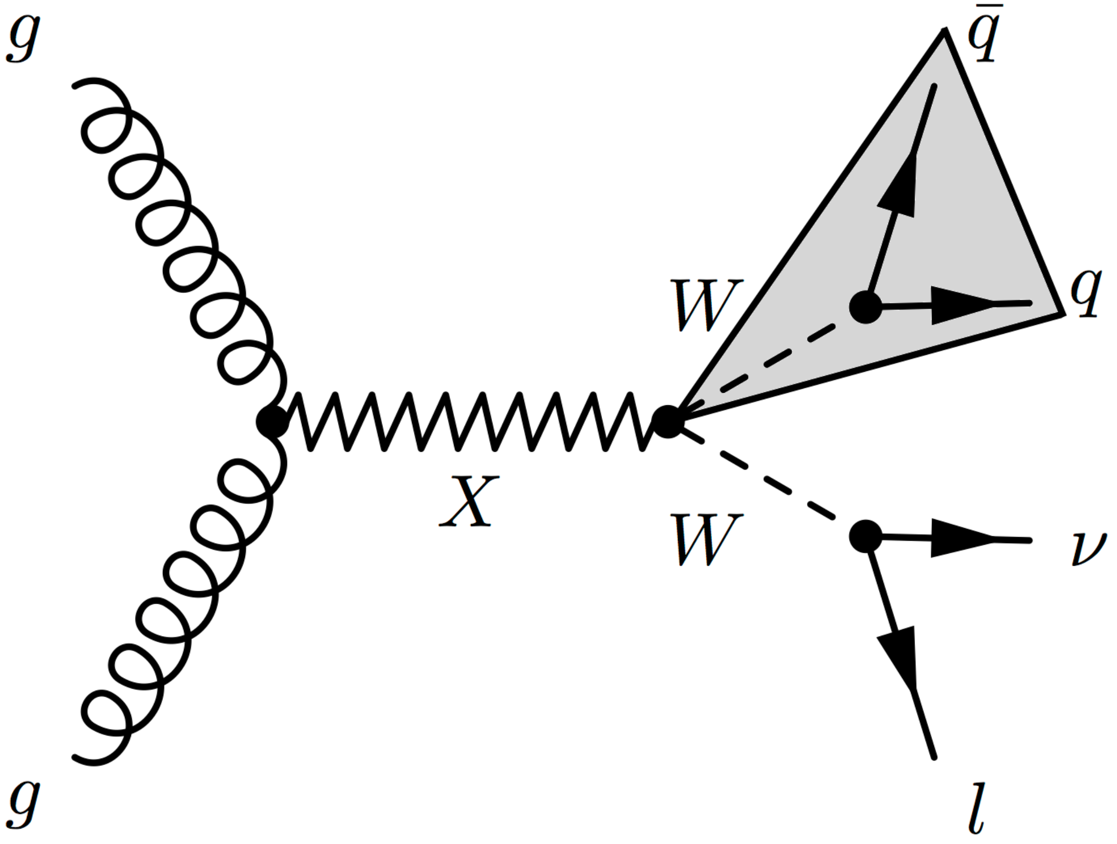


**all-hadronic 13 TeV analyses most sensitive**

# Boson reconstruction

# Reconstructing heavy resonances

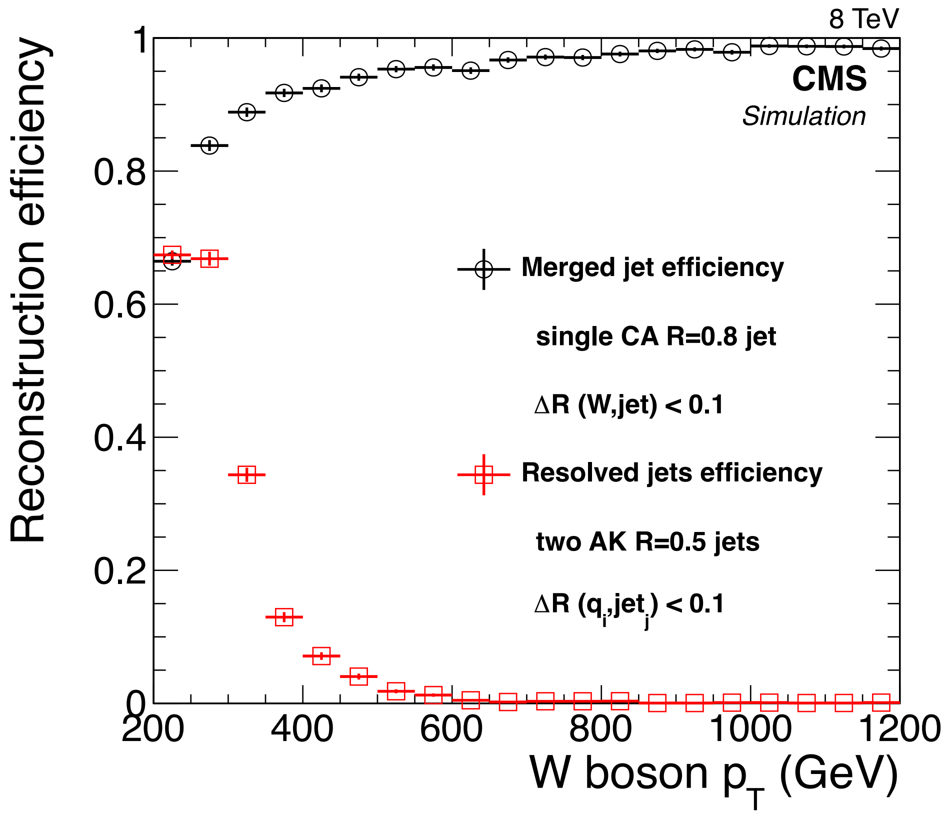
- For high mass resonances, bosons will be very energetic
  ➜ **collimated decay products**
- Need to develop dedicated reconstruction methods
- Hadronic decays of bosons:
  - „**boson-tagging**“
  - exploiting **substructure** of jets
- Leptonic decays:
  - special **isolation** for dileptonic decays
  - dedicated reconstruction algorithms for high-$p_T$ leptons
  - new τ-identification algorithms



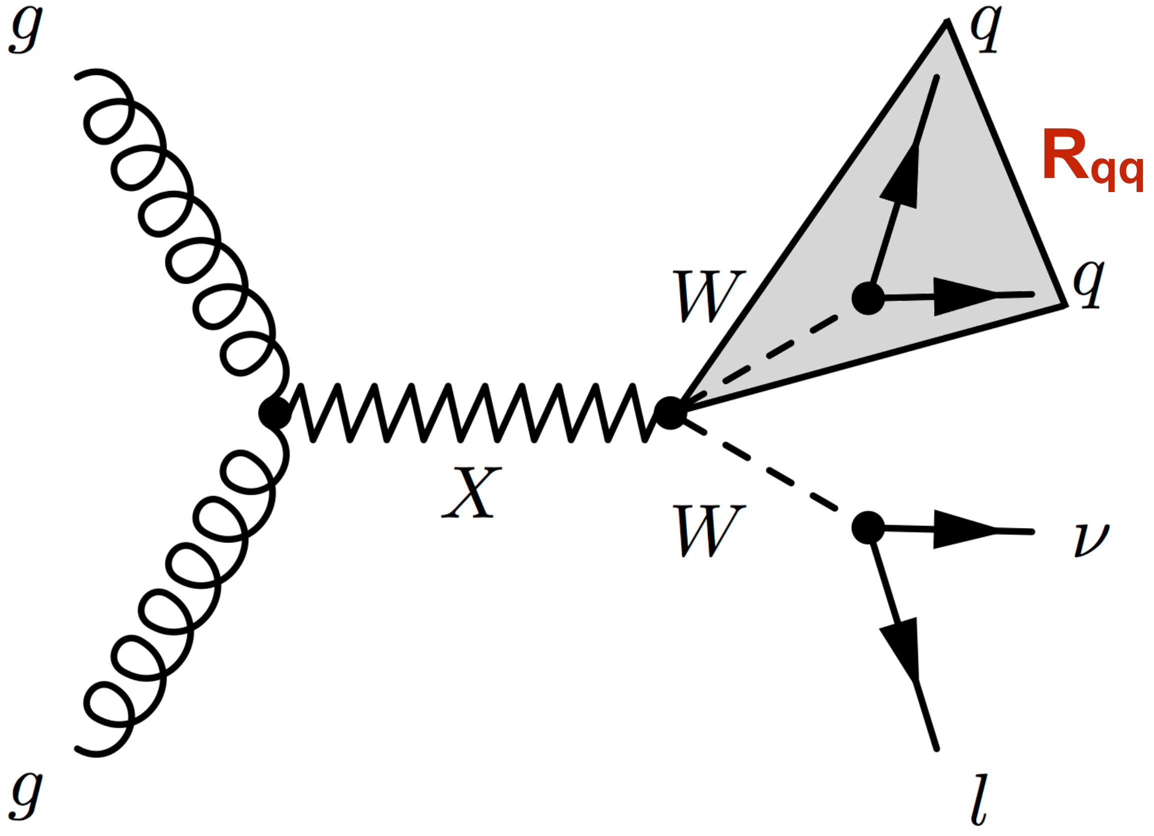
JME-13-006 (JHEP 12 (2014) 017)

# Hadronic boson identification

- At CMS use anti-k$_T$ jet algorithm with R = 0.4
- Already for resonances of **1 TeV** a significant fraction of cases where the boson **decay** is contained **in a single jet**
- Increase jet size to R = 0.8 to contain full decay within **„fat“ jet**



$$R_{qq} = \sqrt{\Delta\eta^2 + \Delta\phi^2}$$



$$R_{qq} \approx 2\frac{M_V}{p_T^V}$$

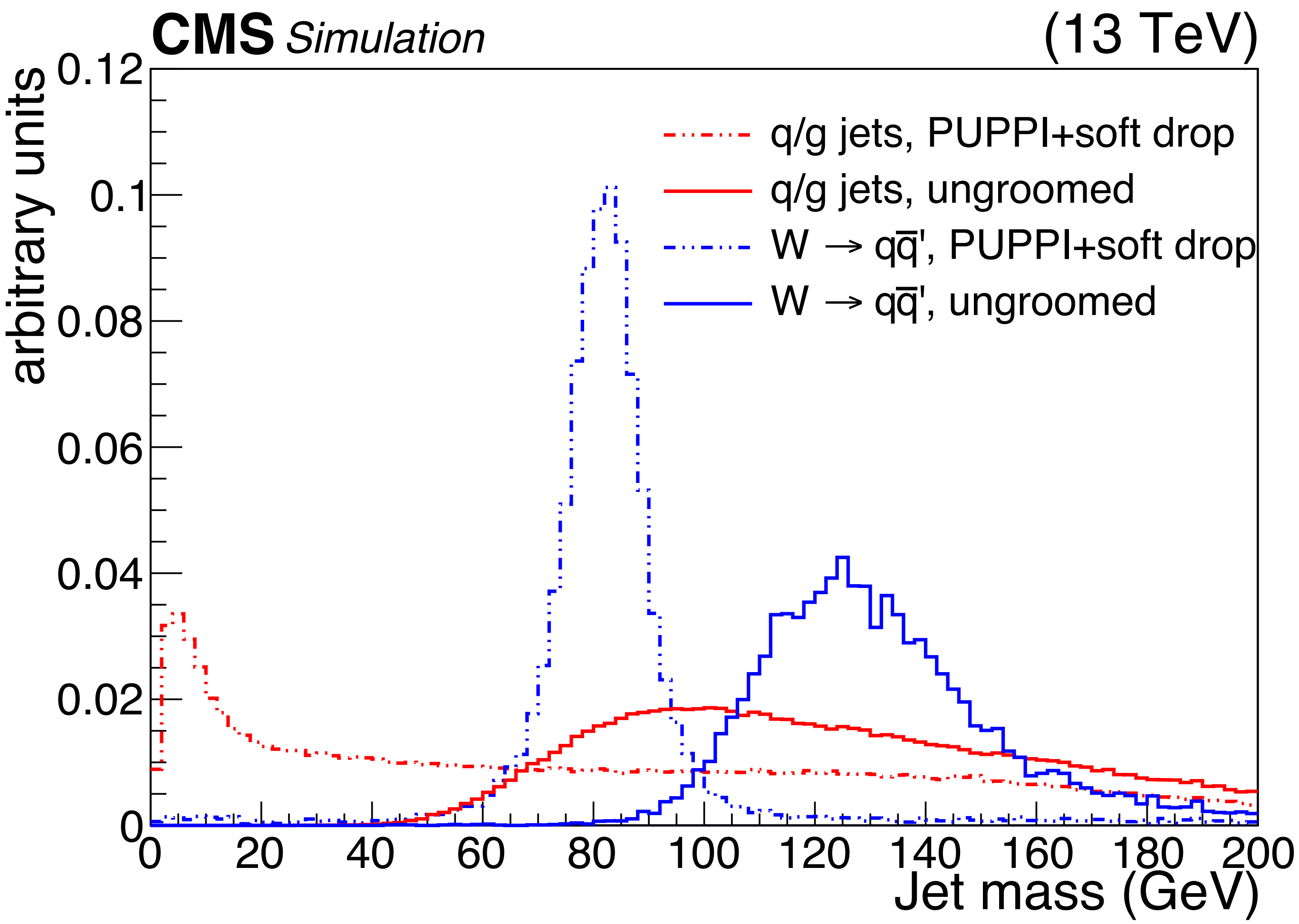
back-of-the-envelope calculation:

**for a resonance of mass 1 TeV the bosons from the decay will have p$_T$~0.4 TeV ➜ ΔR ≈ 0.4**

# Jet grooming

- We know the **masses** of W, Z and Higgs very well ➜ can use them as **constraints**
- However, large number of particles in jet ➜ rather bad resolution
- Jet grooming removes **soft and large angle radiation**
- Strategy:
  - recluster jet using Cambridge-Aachen (CA) jet algorithm
  - iteratively break into two subjets

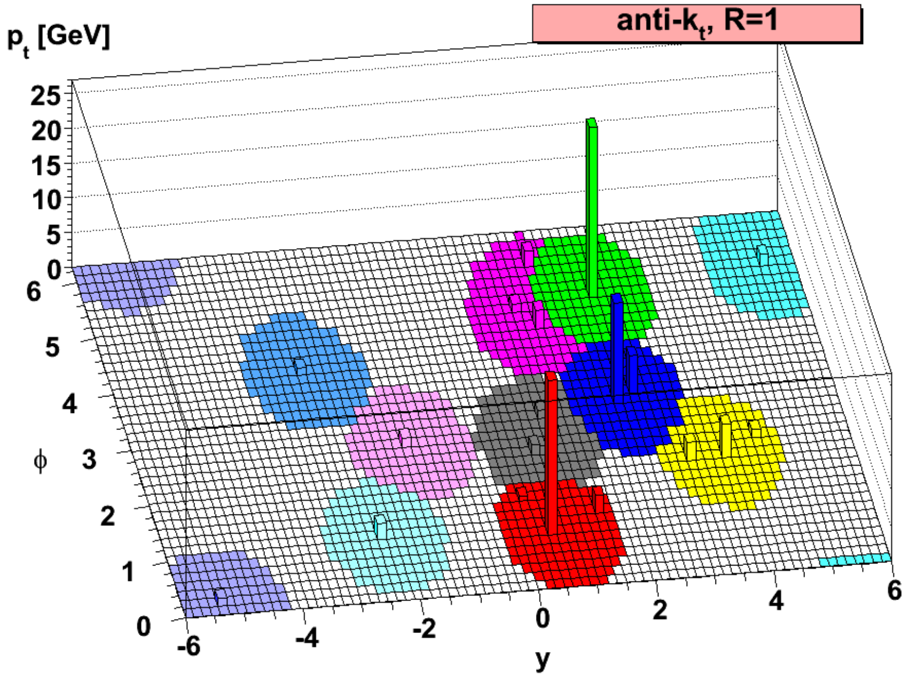
**?**

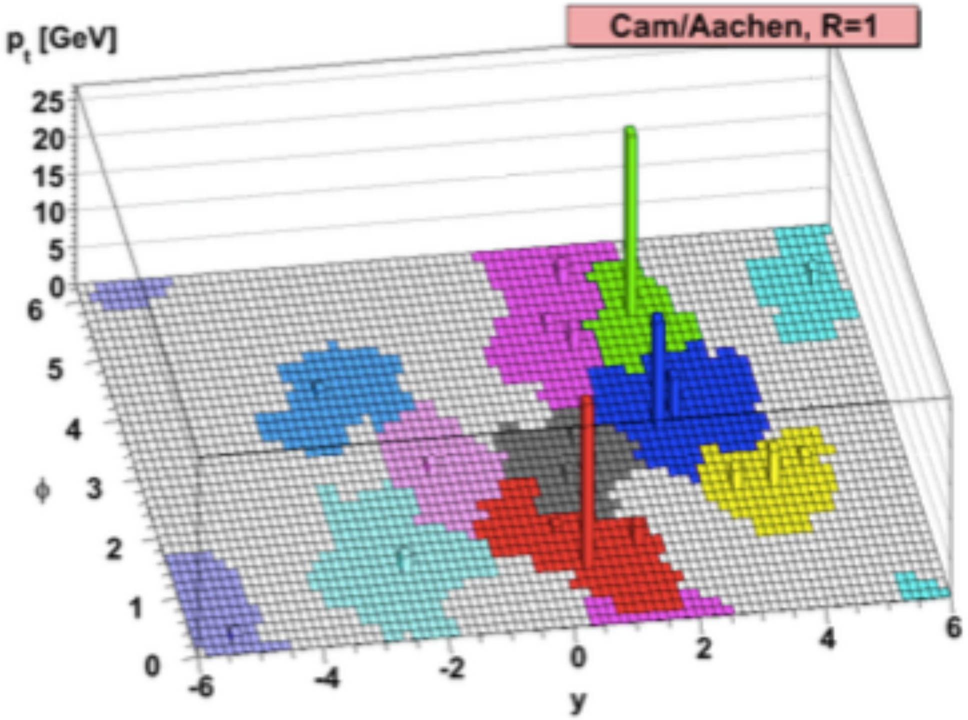


JHEP 0804:063,2008

# Reminder: jet clustering algorithms

- $k_T$-algorithms: **sequential clustering**
- Examine four-vector inputs pairwise and construct jets hierarchically
- **anti-$k_T$**: preferentially merge constituents with **high $p_T$** with respect to their nearest neighbours first
  - undoing merging yields one high- and one low-$p_T$ subjet
- **Cambridge-Aachen**: no $p_T$-weighting, merge based on **spatial separation** only ➜ undoing clustering yields more **$p_T$-symmetric subjets**





# Jet grooming

- We know the **masses** of W, Z and Higgs very well ➜ can use them as **constraints**
- However, large number of particles in jet ➜ rather bad resolution
- Jet grooming removes **soft and large angle radiation**
- Strategy:
  - recluster jet using Cambridge-Aachen (CA) jet algorithm
  - iteratively break into two subjets
  - remove softer contribution (and continue with harder one) if:

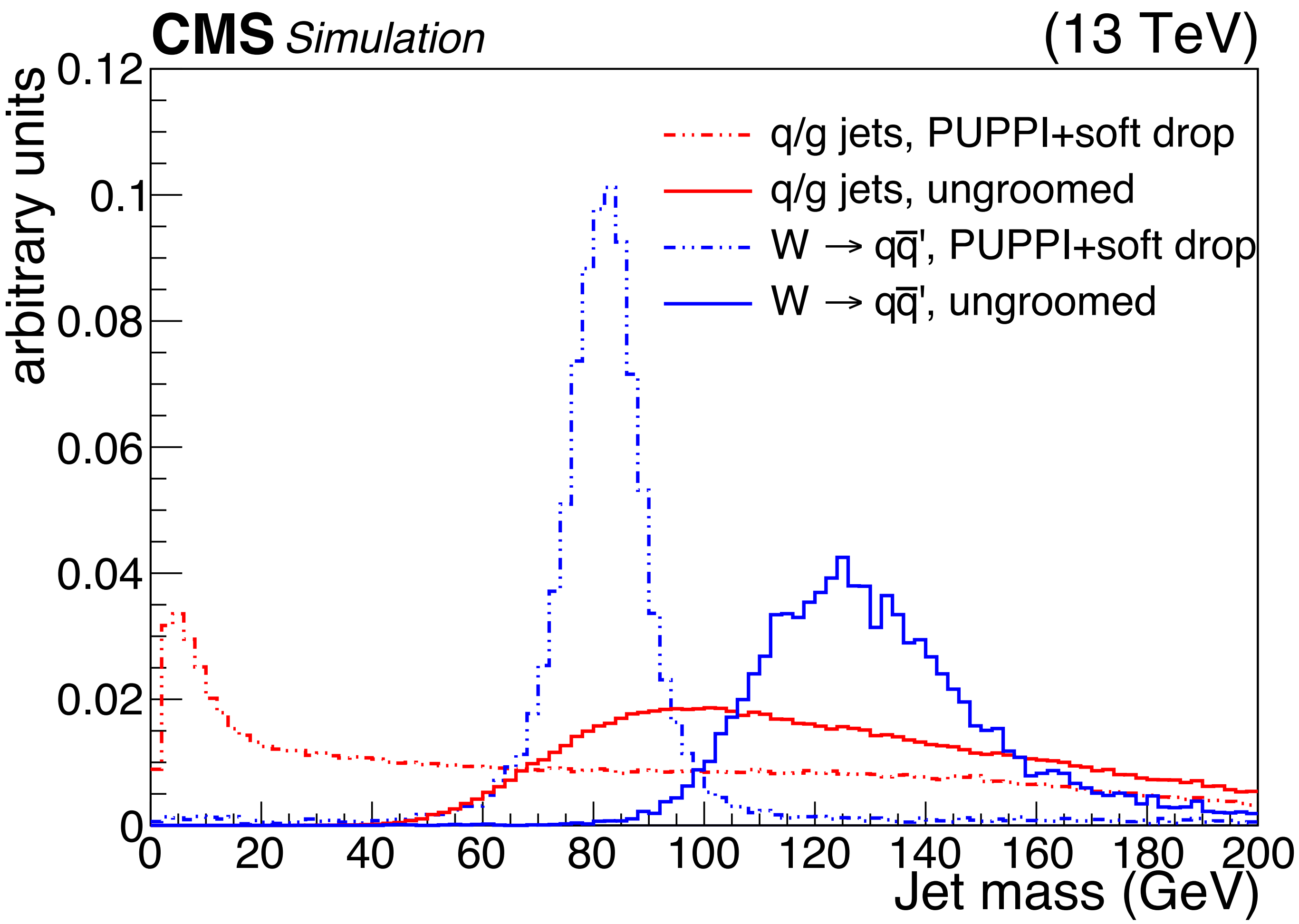
**soft-drop condition**

$$\frac{\min(p_{T1}, p_{T2})}{p_{T1} + p_{T2}} < 0.1$$

modified mass-drop algorithm (mMDT)

  - stop otherwise
- Cut on mass window (~±10 GeV)

**before grooming, pileup removal is performed (PUPPI)**



**use of soft-drop algorithm new w.r.t. 2015 — previously used pruning ➜ perturbatively robust**

JHEP 1103:015,2011

# N-subjettiness

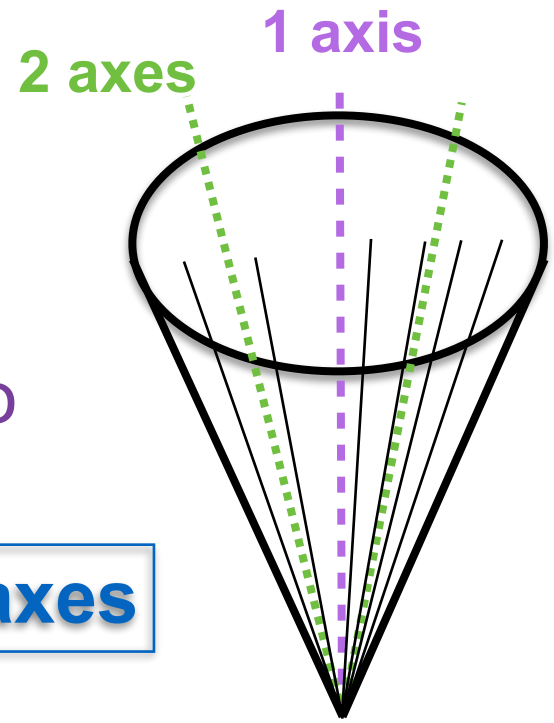
- For **boson-tagging**: want to quantify how **2-subjetty** a jet is
- ➜ To what extent is energy flow aligned along 2 momentum directions (N=2)?

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

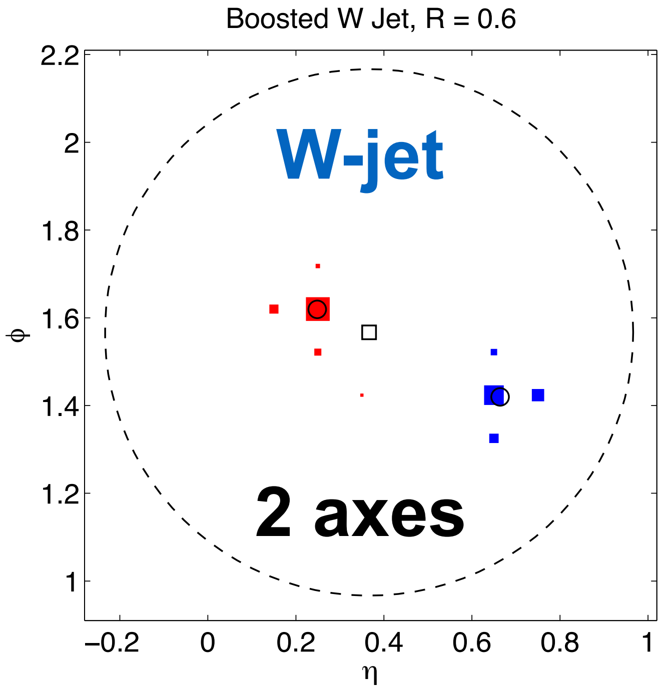
normalisation

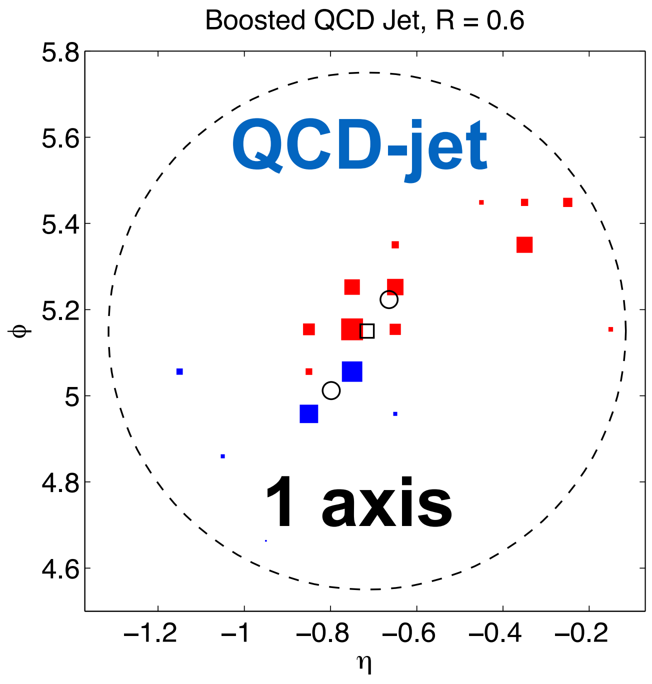
sum over particles

minimise distance to candidate subjets



**low values of τN ➜ compatibility with the hypothesis of N axes**

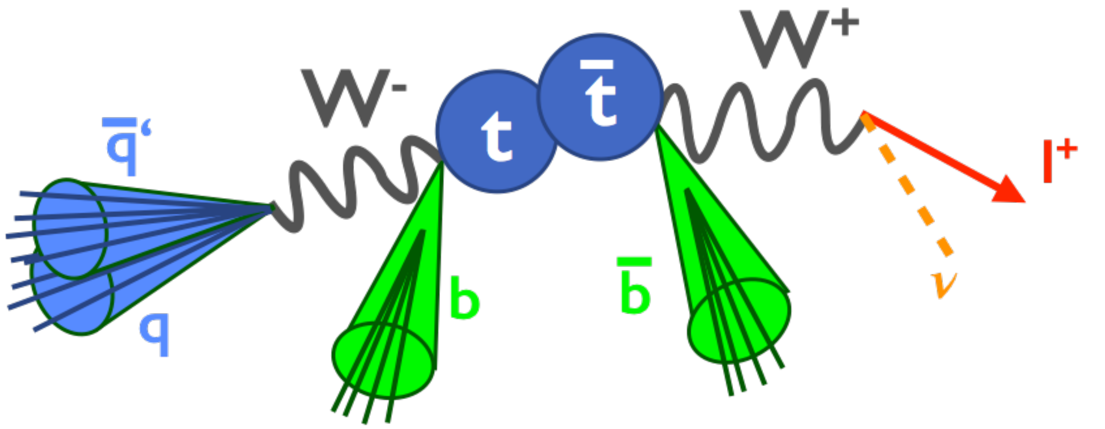

Boosted W Jet, R = 0.6

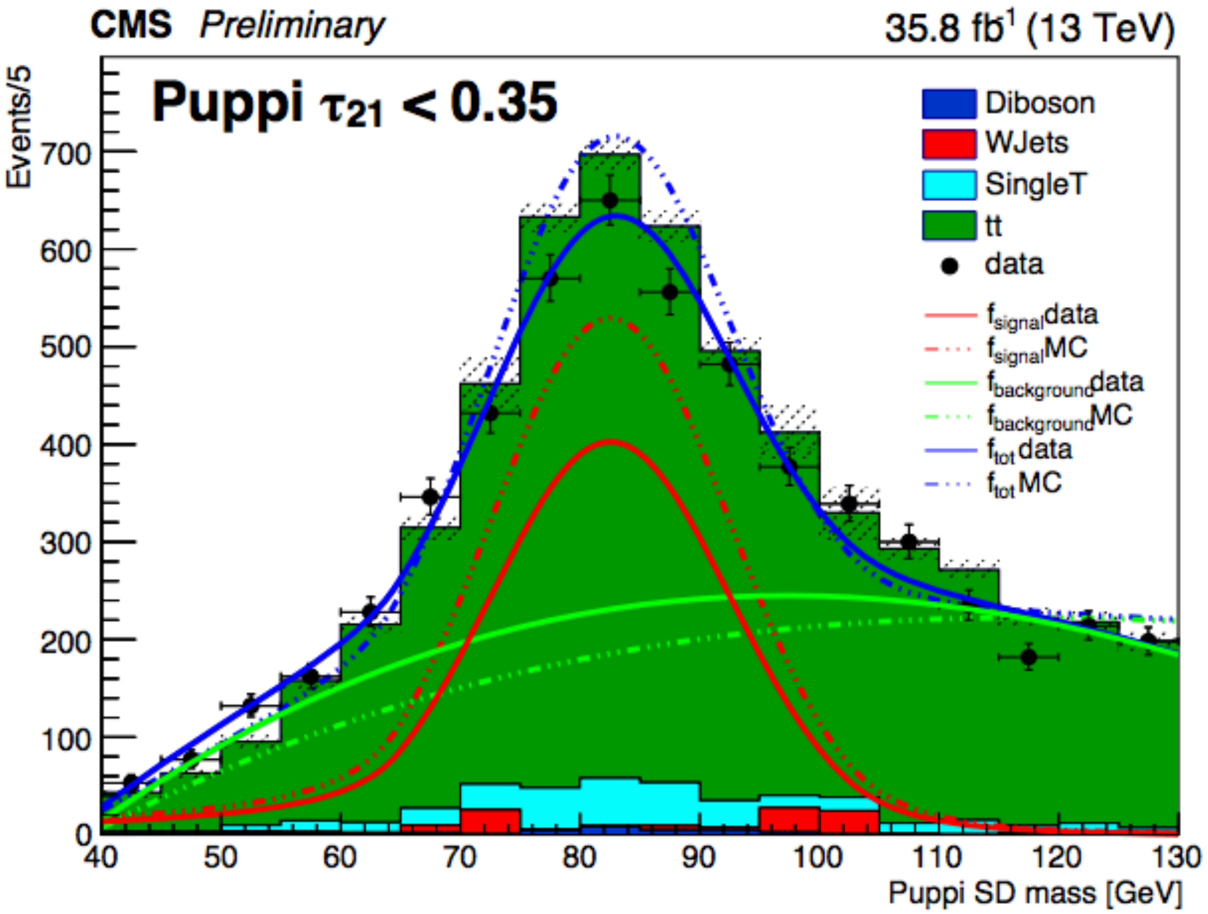

Boosted QCD Jet, R = 0.6

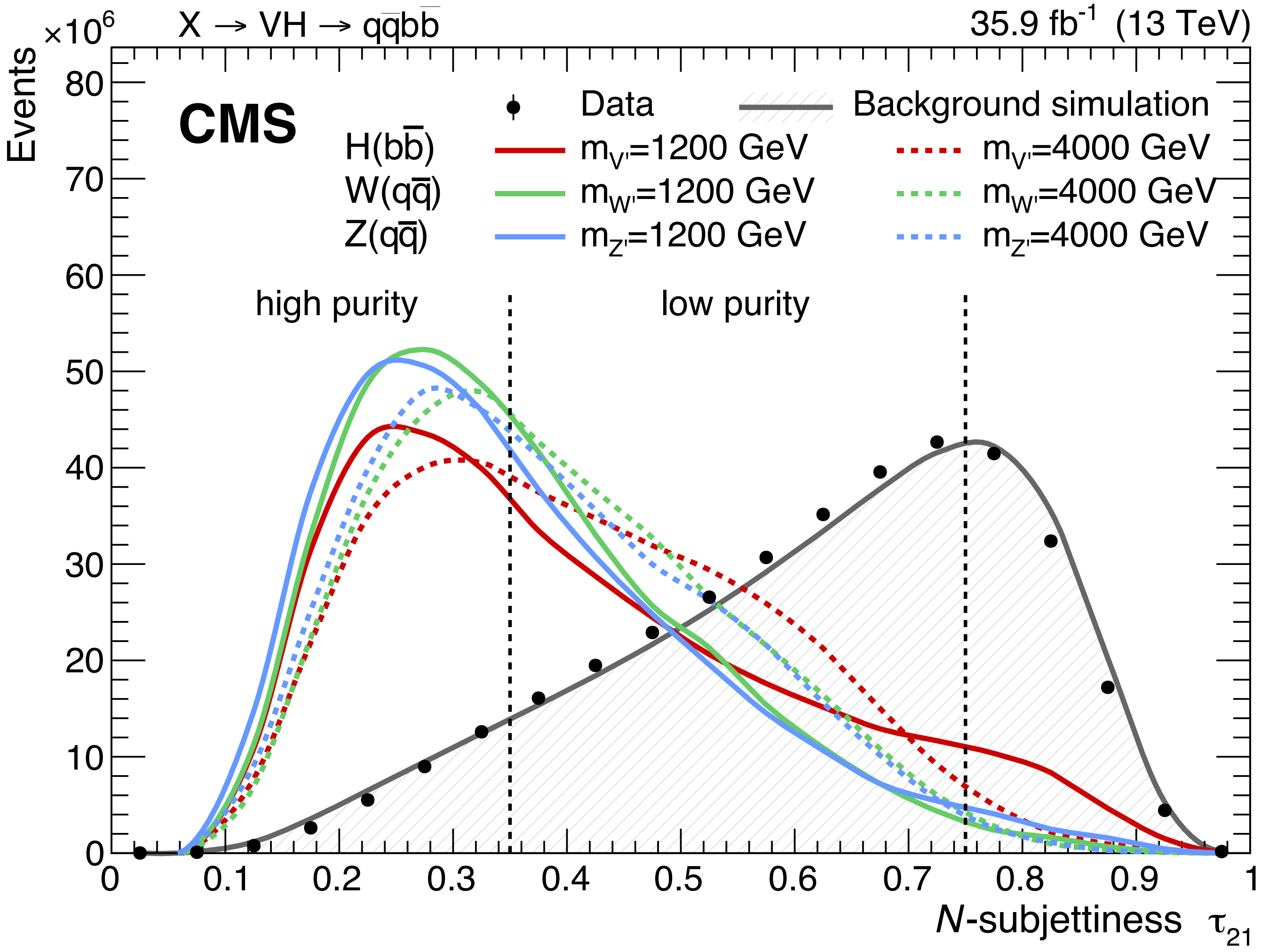
# N-subjettiness ratio

B2G-17-002 (Eur. Phys. J. C 77 (2017) 636)
CMS DP-2017/026

- Bare $\tau_N$ has very little discrimination power
- Take **ratio $\tau_2/\tau_1$** instead
- Note: rather **complicated variable**, difficult to model ➜ need to validate in data
- Clean sample of W-jets: **top-antitop quark pairs** used for calibration (W-jet $p_T$ ~ 200 GeV)

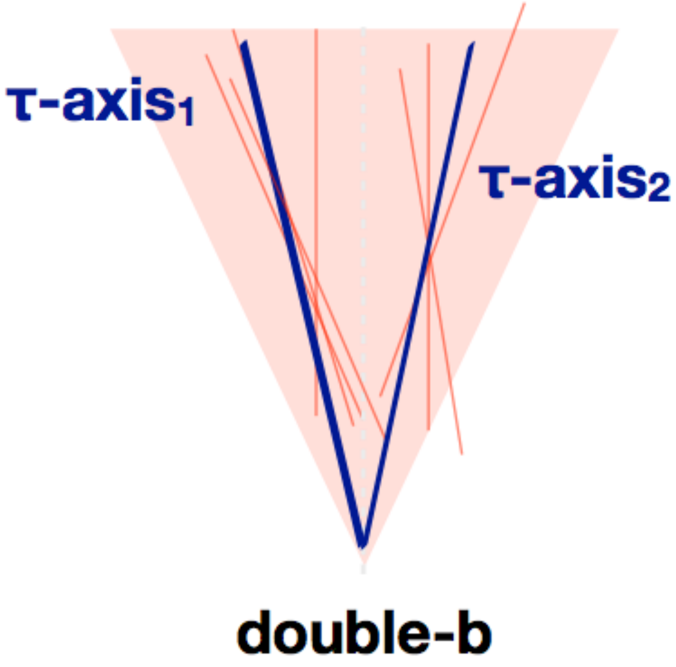


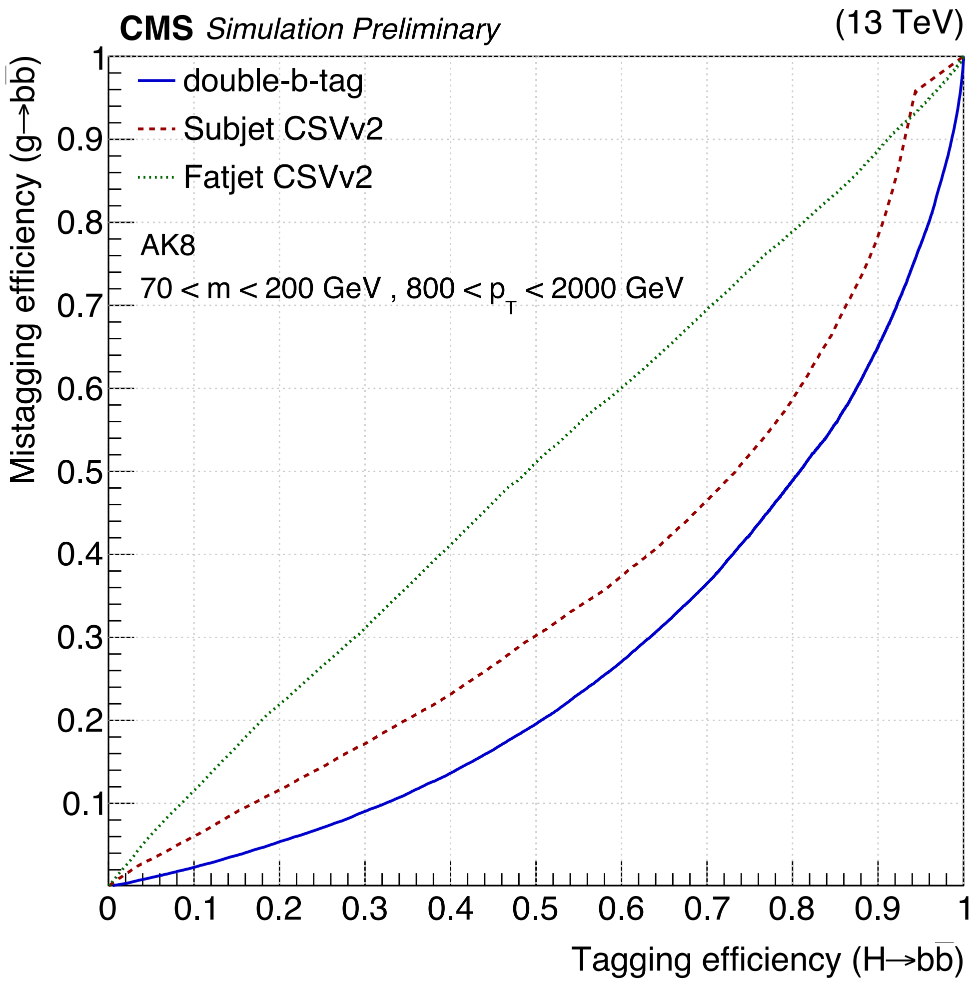




# Higgs➜bb tagging

- Higgs has higher mass than W/Z bosons ➜ $\tau_2/\tau_1$ less important, **exploit b-jet content** instead
- Previously, two different strategies:
  - identify b-subjets
  - tag fat jet
- Already **50% lower mis-tagging rate** than W-/Z-tagging
- Run-2: dedicated Higgs-tagger — **double-b tagger** (MVA-based):
  - inputs based on observables from secondary vertex and tracks associated to each τ-axis (27 total)
  - significantly better background rejection
- For resolved Higgs decays: new DeepCSV discriminator





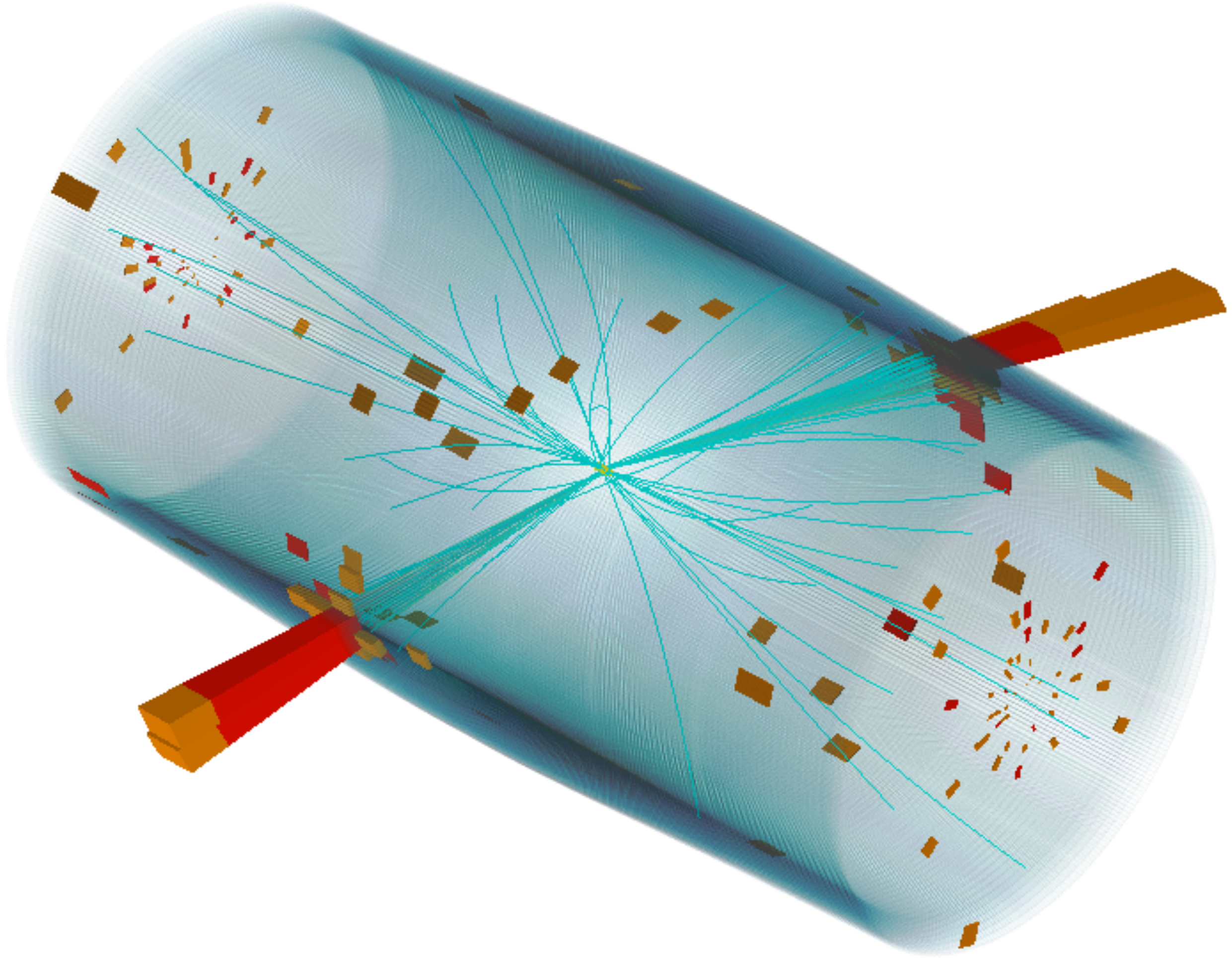
# All-hadronic VV/Vh final states

# All-hadronic final states

- Boson branching fractions to quarks very large:
  - W/Z ➜ quarks ~69%
  - Higgs ➜ bb ~57%
- High acceptance
  - large reach in resonance mass
- For heavy resonances: dijet final state
- Challenge: background estimation

**Candidate ZZ event**

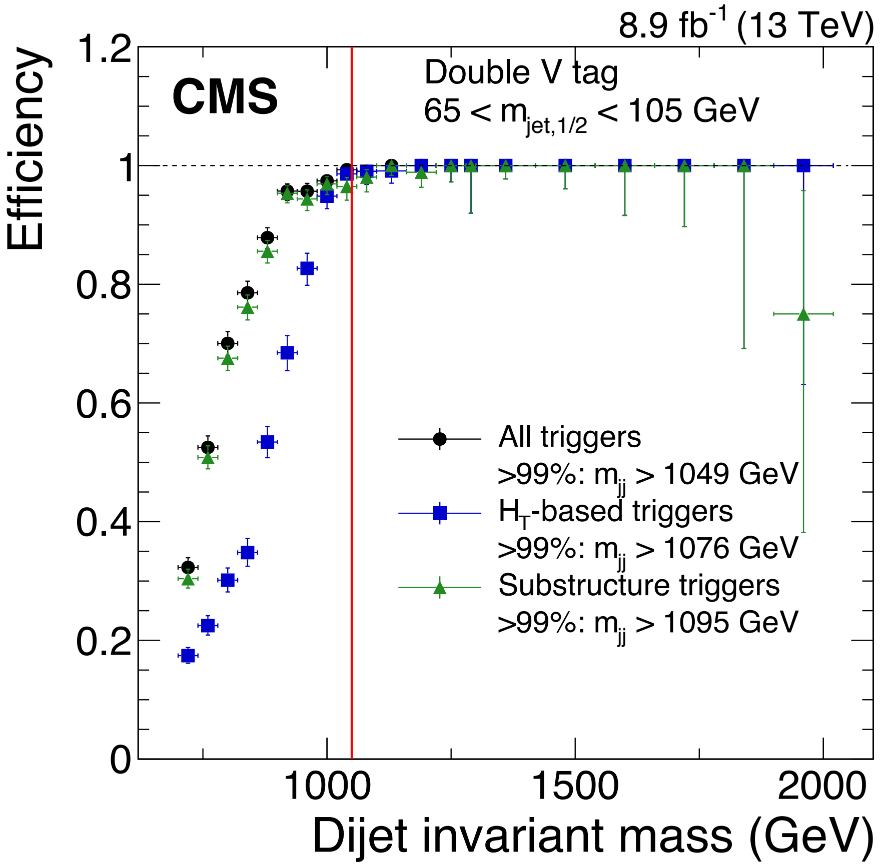
Dijet mass: 3.2 TeV

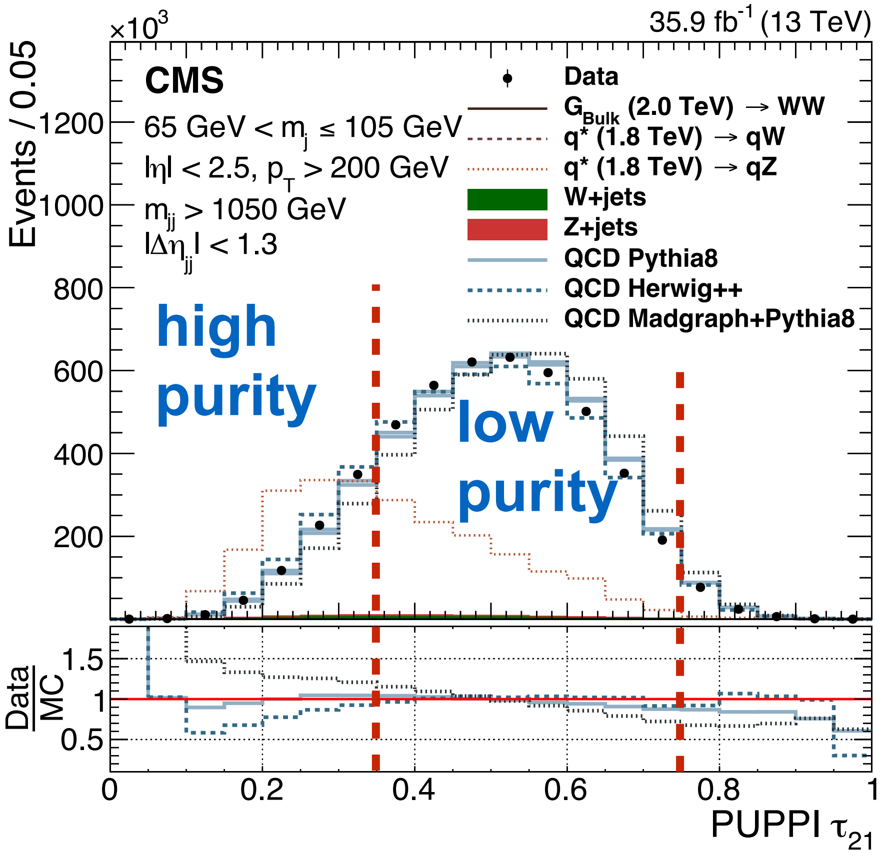


B2G-17-001 (submitted to Phys. Rev. D)

# All-hadronic VV analysis

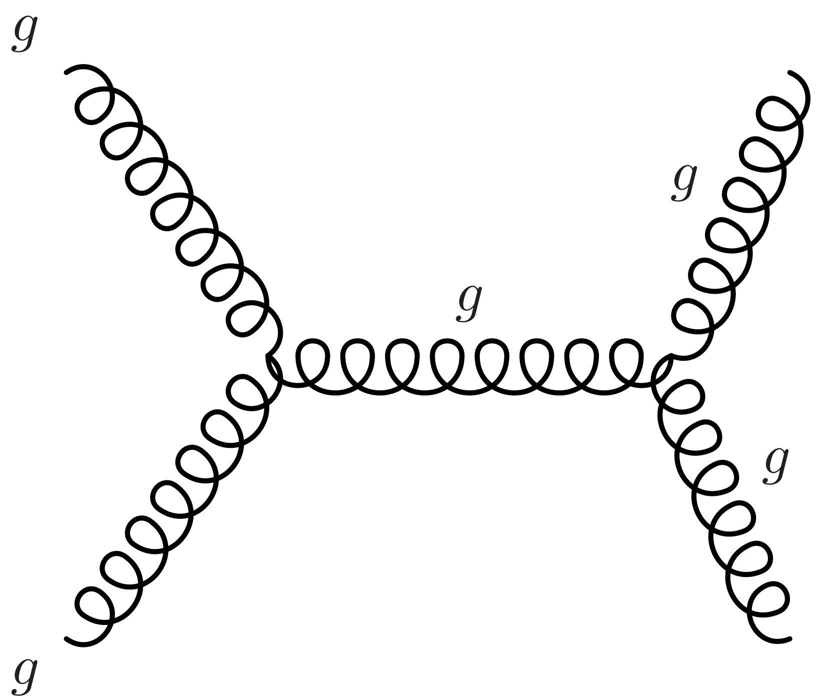
- Dijet final state: WW, WZ, ZZ (=VV)
- High trigger thresholds
  - $H_T$-based (scalar $p_T$ sum of all objects in the event)
  - high-$p_T$ jet with substructure
- **Trigger** at ~100% efficiency for **$m_{JJ}$ > 1.05 TeV**
  - apply cut on reconstructed dijet system
- Define different **$\tau_2/\tau_1$ regions**:
  - high purity to suppress background
  - low purity to recover signal efficiency at high masses
- **Split W and Z** samples based on soft-drop jet mass (65-85, 85-105 GeV)

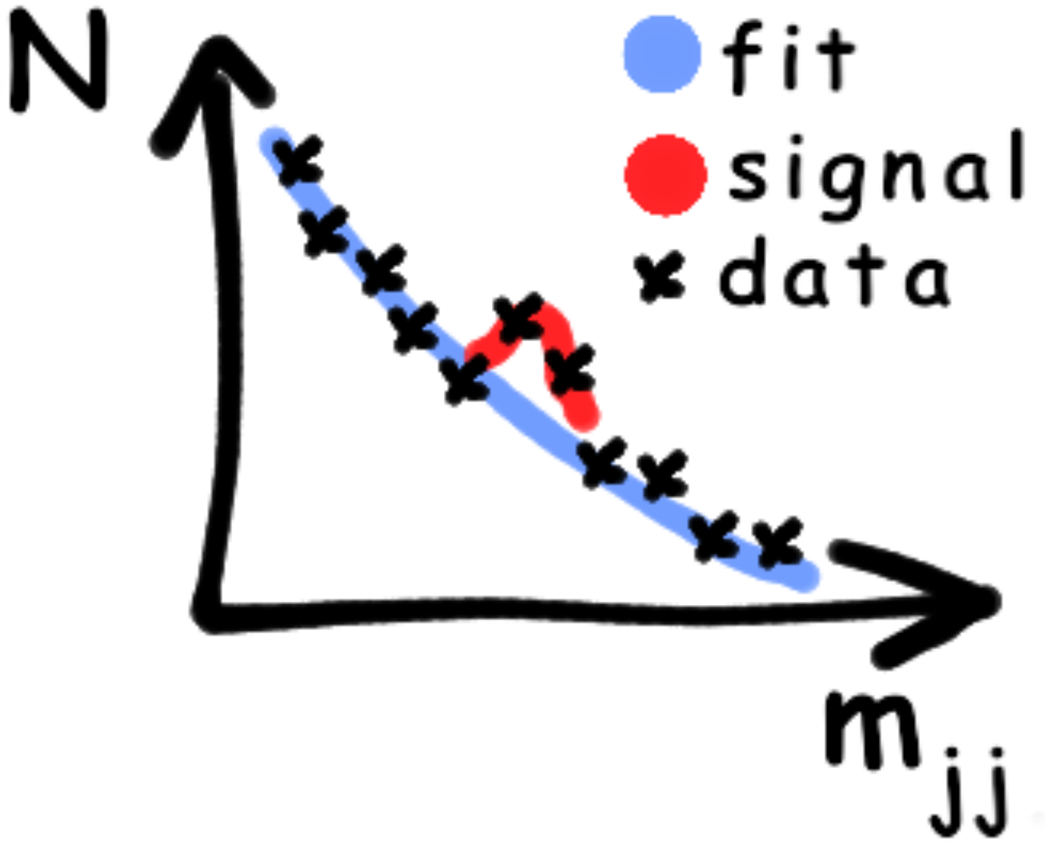




# VV background estimation (1)

- After selection, still dominated by **QCD multijet** events
- Difficult to obtain sufficient number of simulated events
- Need a data-driven approach
- Exponentially falling spectrum
  - since we are in the trigger efficiency plateau
- Can use **fit function** and perform a so-called **bump hunt**





# VV background estimation (2)

B2G-17-001 (submitted to Phys. Rev. D)
B2G-17-002 (Eur. Phys. J. C 77 (2017) 636)

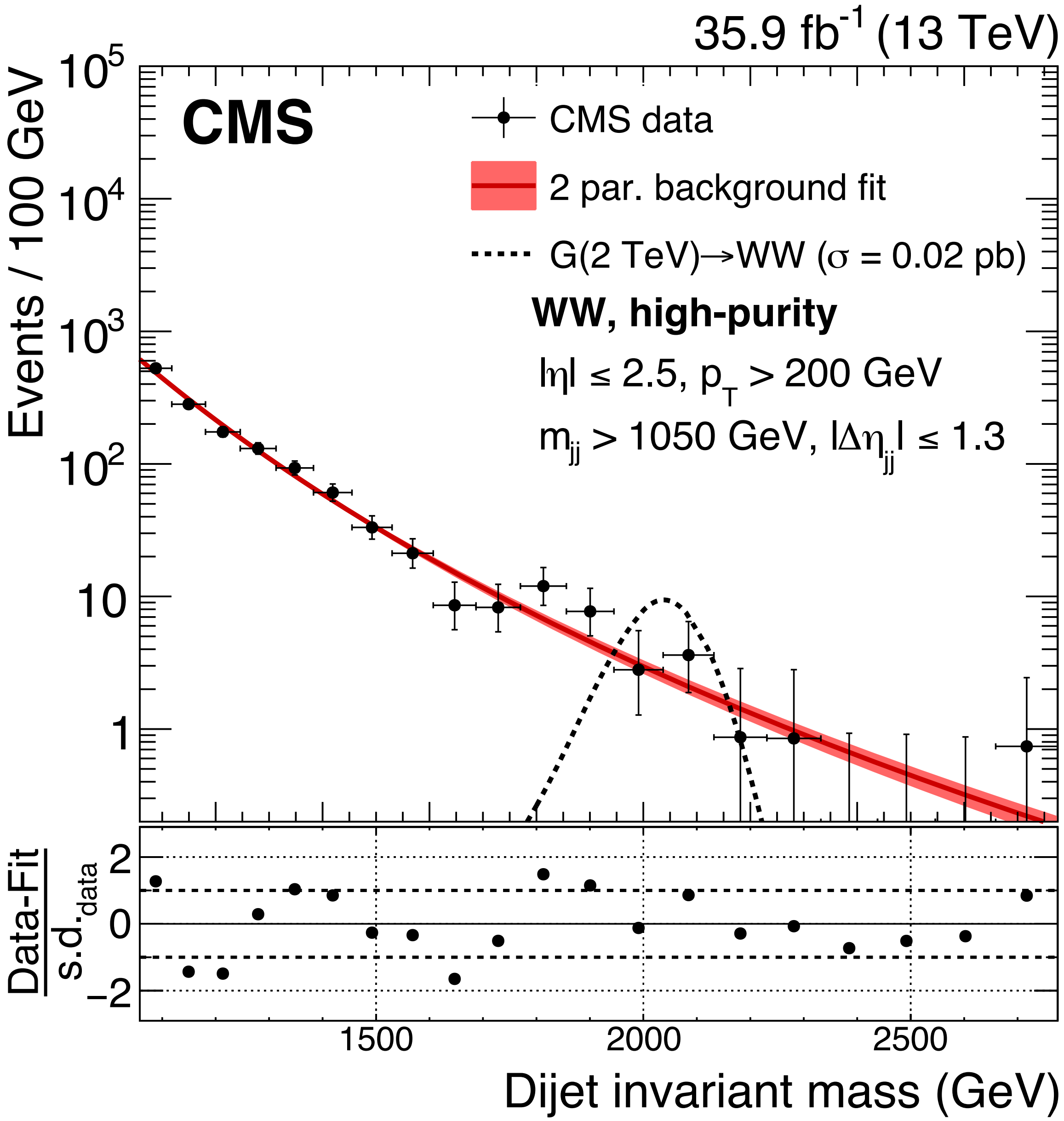
- Naively, fitting the $m_{JJ}$ spectrum could swallow signal
- Also, need to **avoid claiming false discovery** (in particular in tail)
- Fit function:

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

2 parameters 3 parameters

… N parameters

- Number of free parameters determined by **F-test**:
  - check if quality of fit improves by > 10% confidence level
  - if not, stick with current fit function
- **Extensive bias tests** conducted
- Combined **signal+background fit** performed



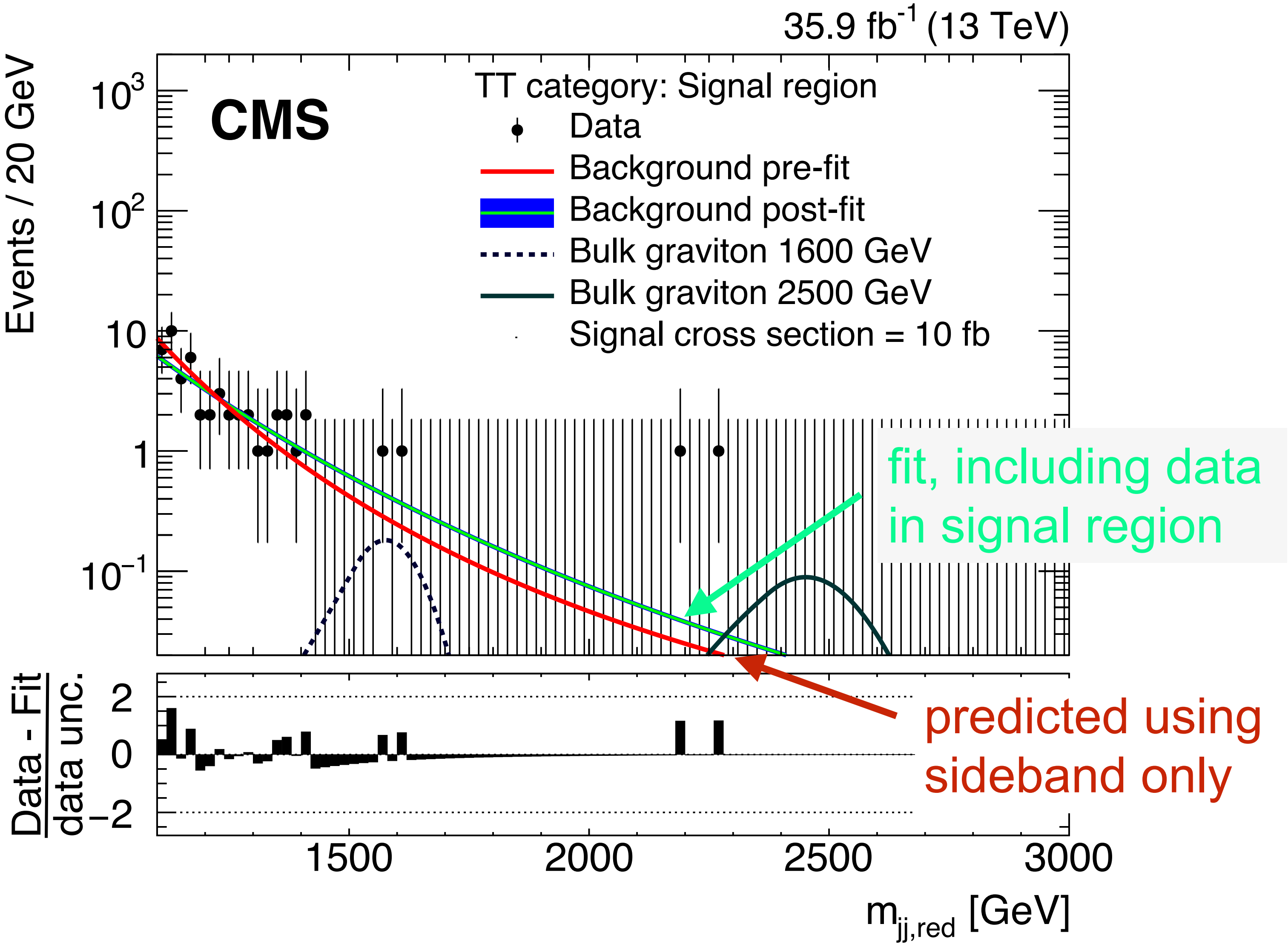
**spin-1 and -2 interpretation**

**see also Vh analysis (B2G-17-002)**

# Di-Higgs (hh) analyses

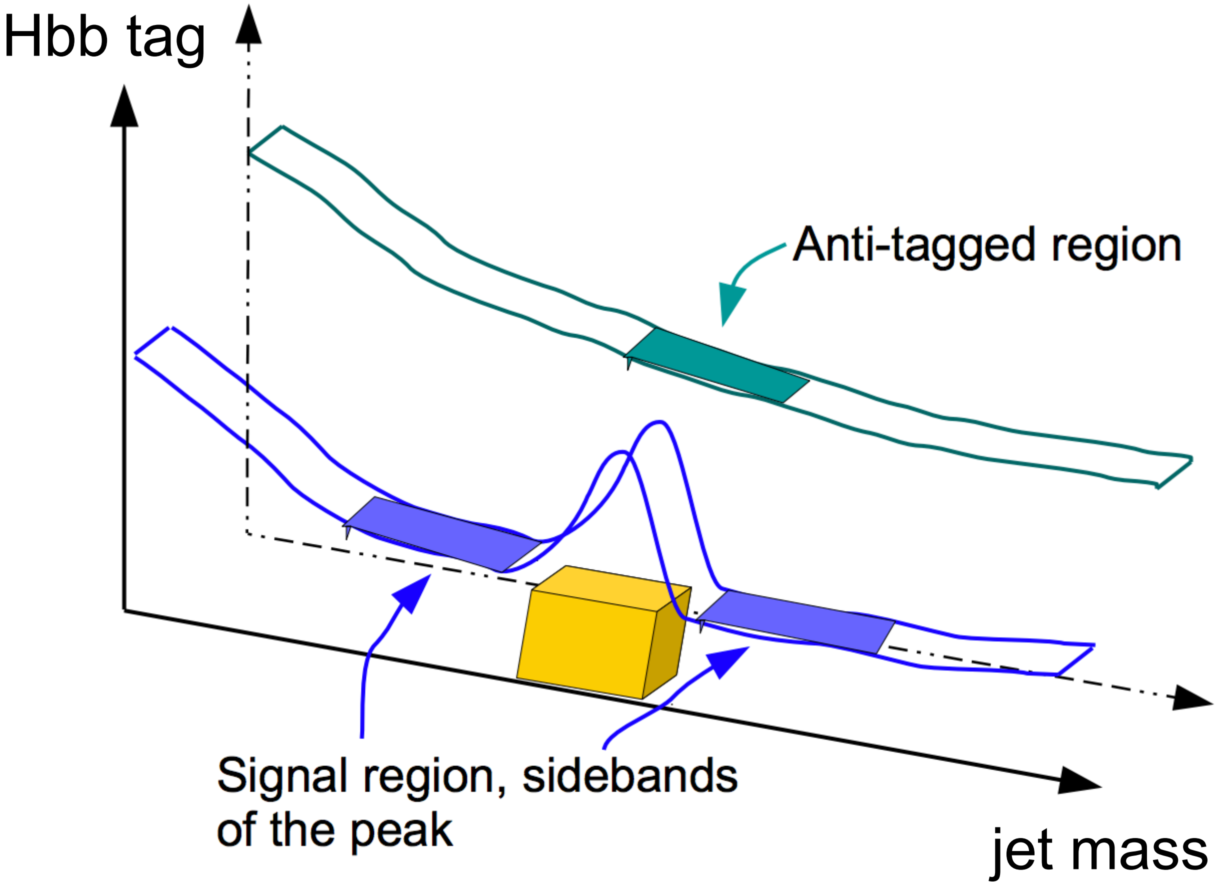
# hh analysis — high mass

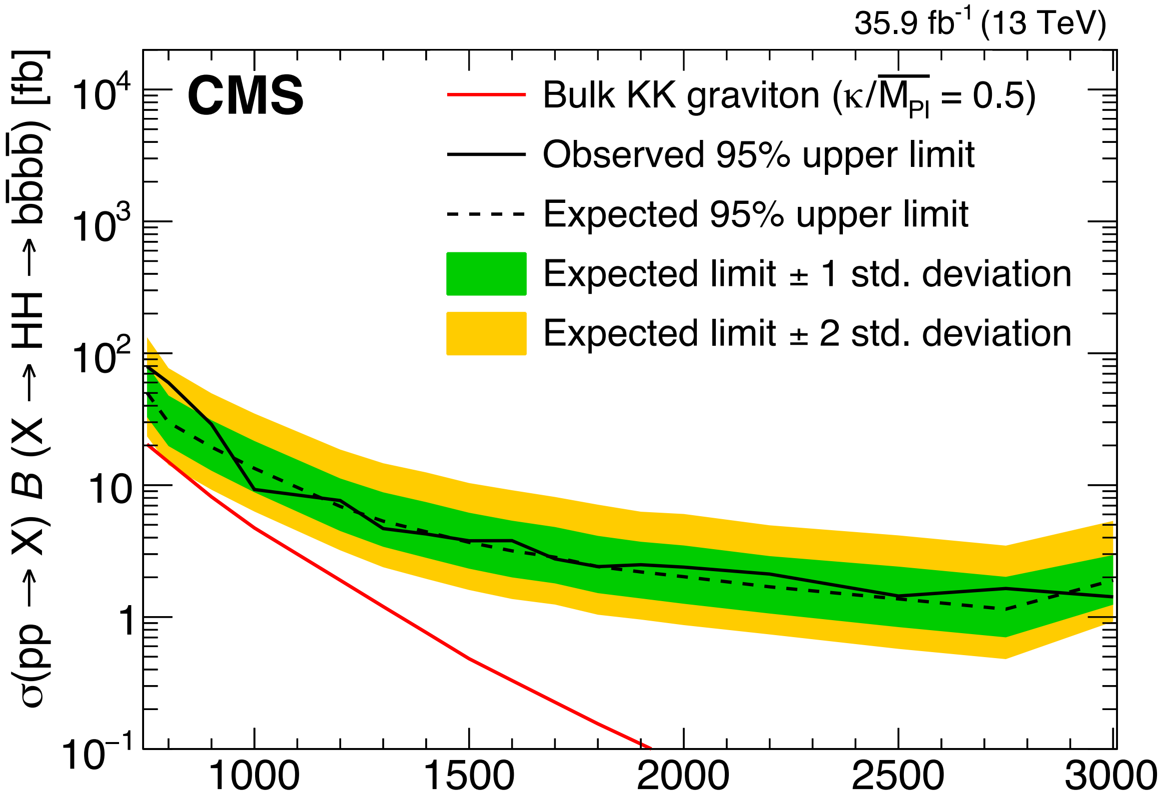
- Dijet final state: hh➜2 bb-jets
- New background estimation technique: Alphabet-assisted bump hunt
  - define anti-Hbb tagged region
  - parametrise anti-tagged to tagged ratio to constrain background from sidebands
  - perform dijet fit



fit, including data in signal region

predicted using sideband only



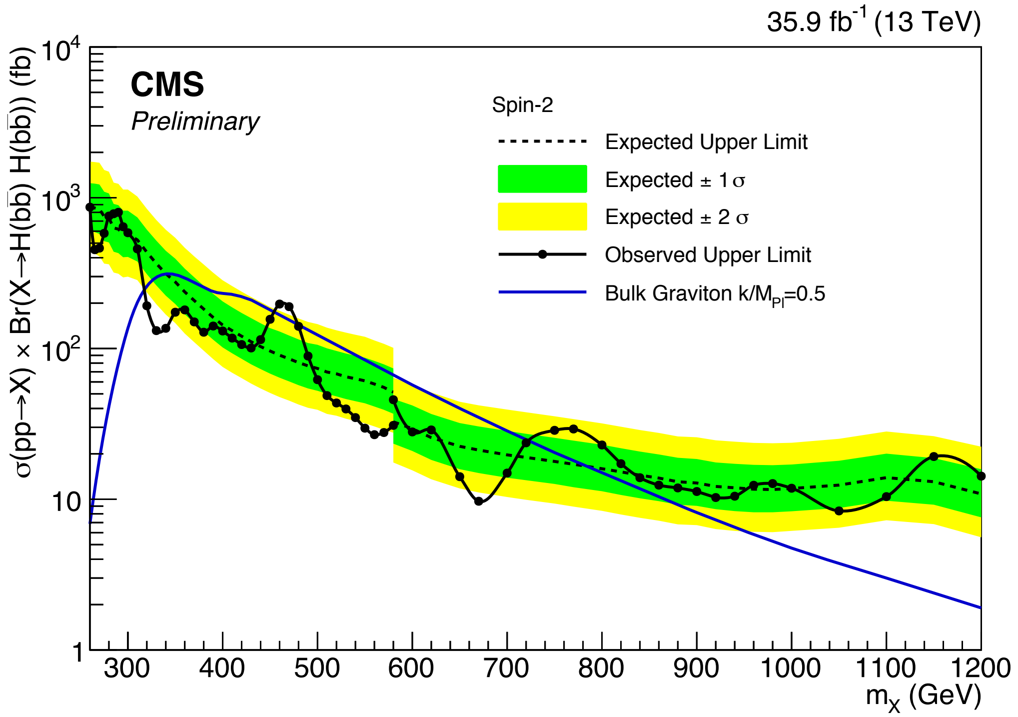


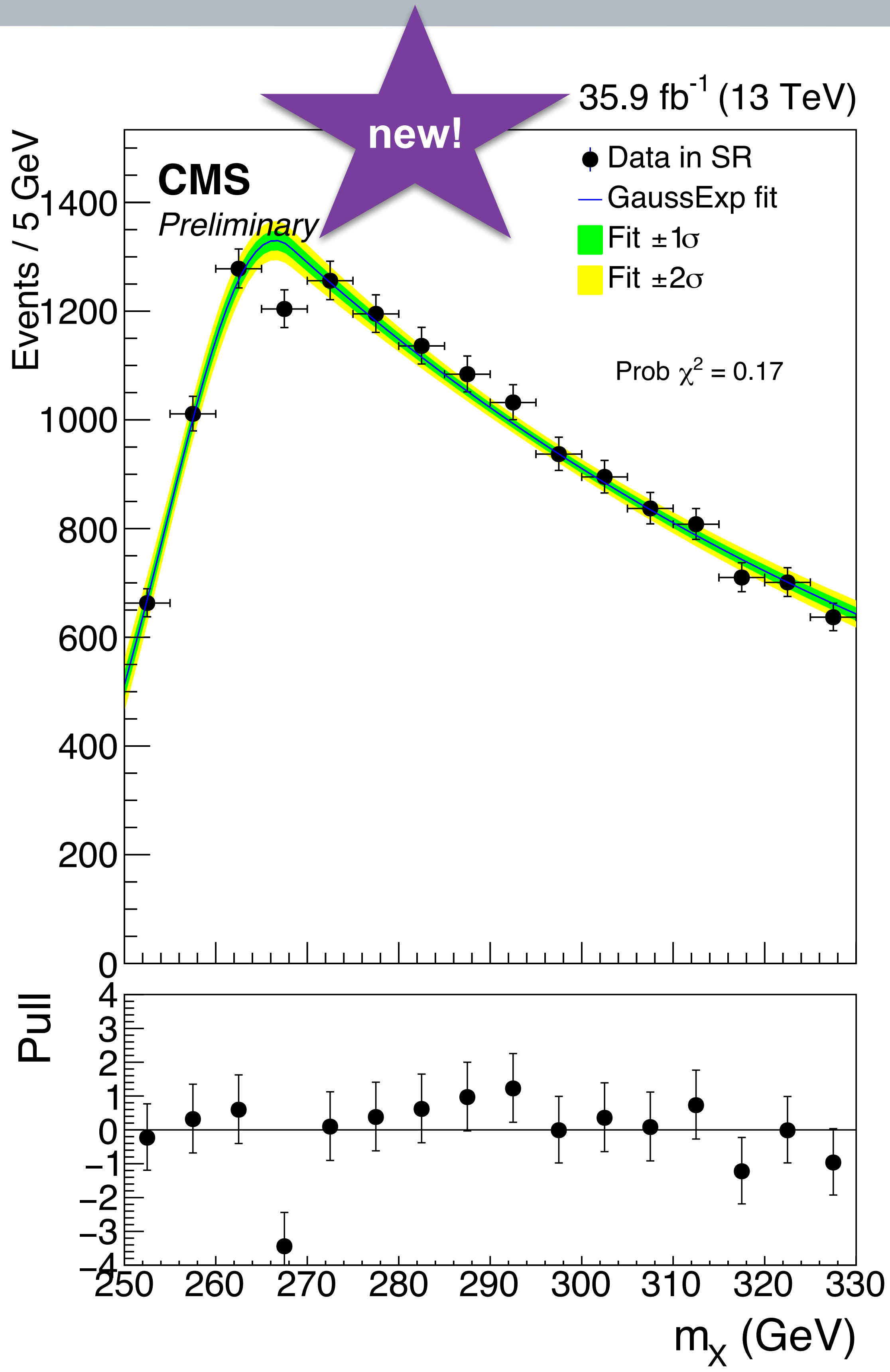
**spin-0 and -2 interpretation**

HIG-17-009

# hh➜4b analysis — low/medium mass

- Can go lower in mass
  - go from boosted dijet final state to (semi-) resolved final state
- Select 4 anti-k$_T$ (R=0.4) jets
  - low mass region: 260-620 GeV
  - medium mass region: 550-1200 GeV
  - profits from DeepCSV b-tagging algorithm
- Also uses parametric background estimation

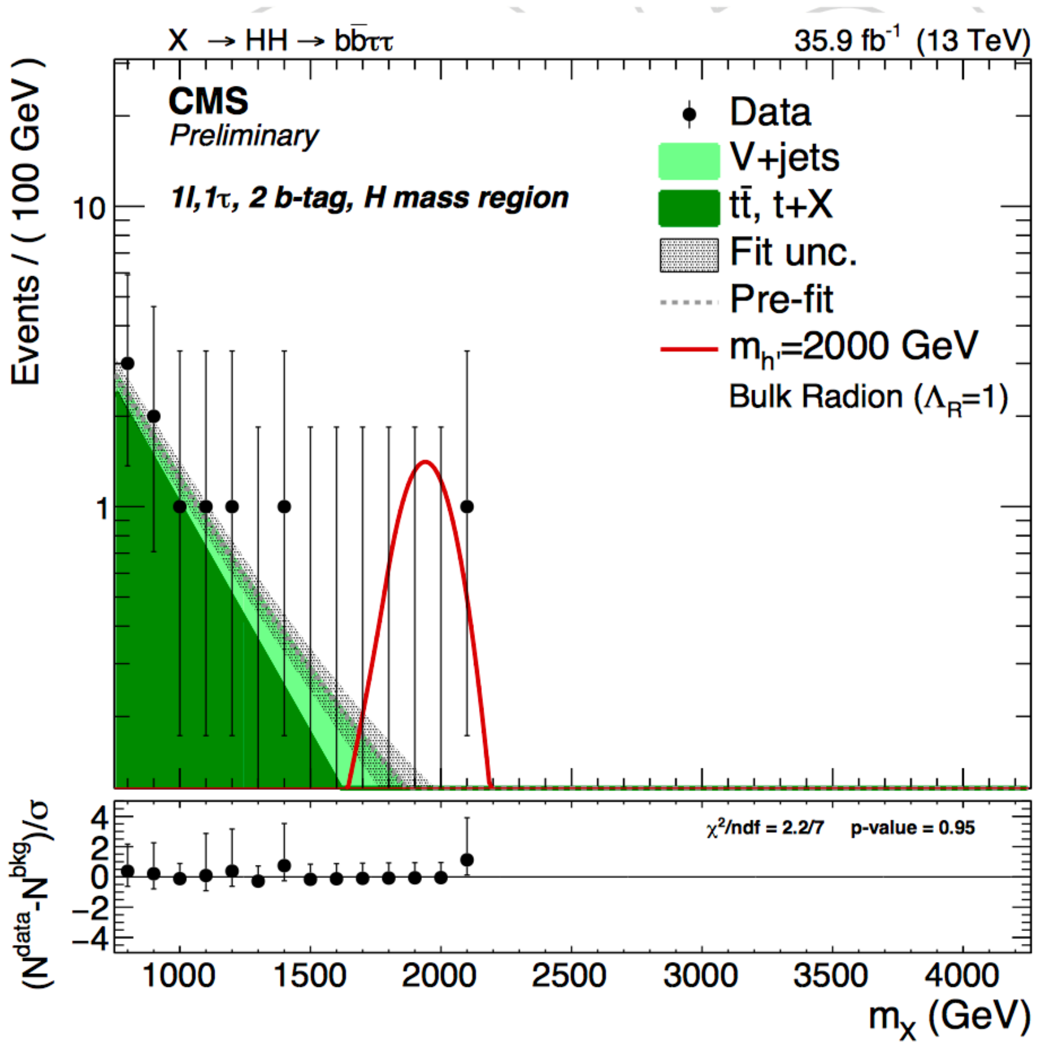


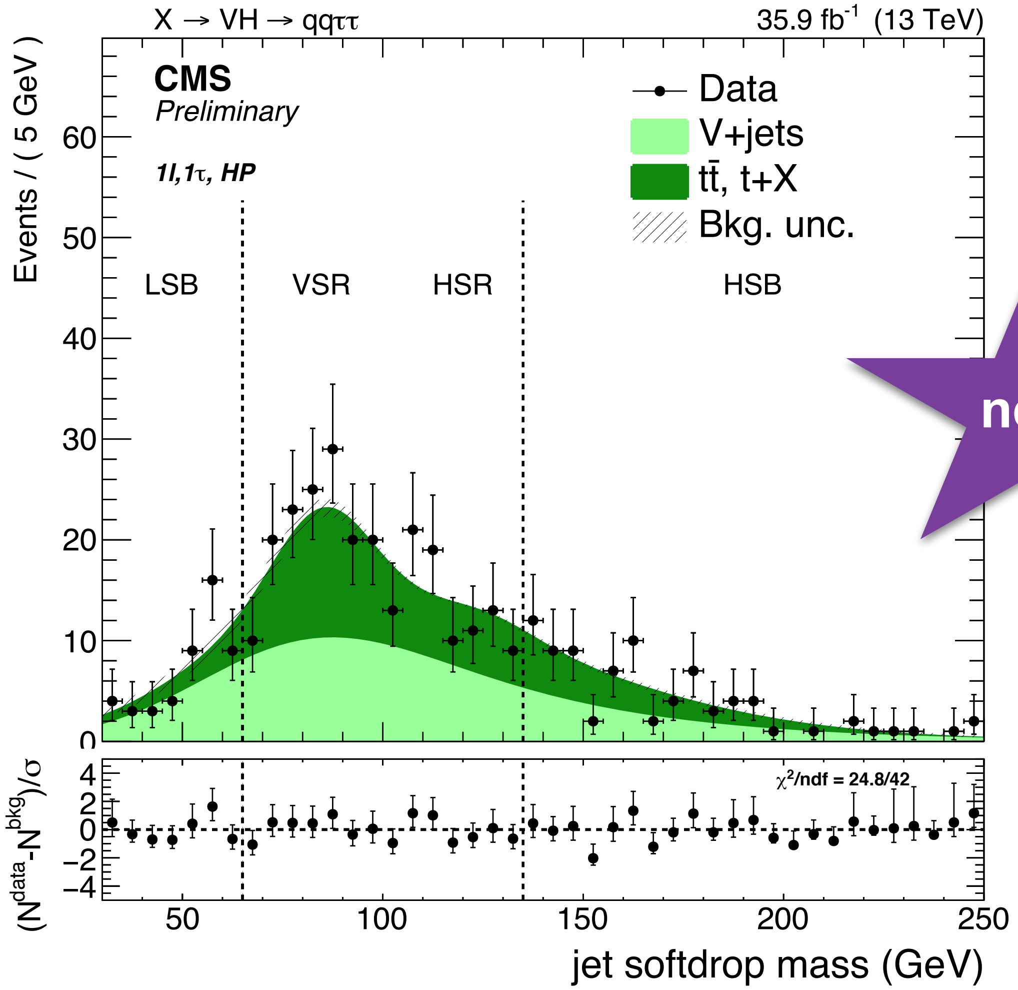


**spin-0 and -2 interpretation**

# More on di-Higgs production

- Large number of di-Higgs analyses in CMS
  - resonant and non-resonant (self-coupling, probing the electroweak symmetry breaking potential) interpretation
  - hh➜4b analyses most sensitive at high invariant hh mass





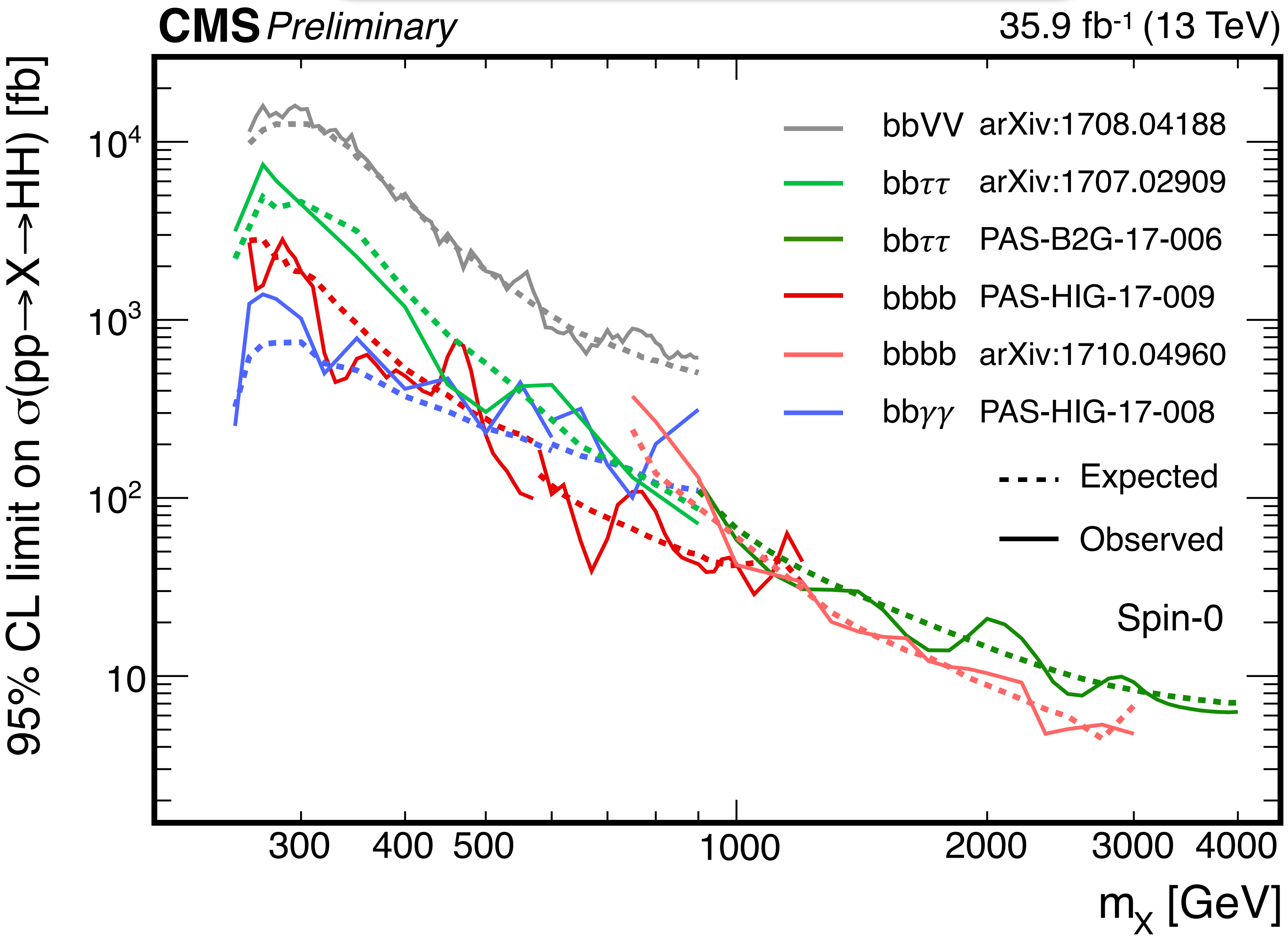
**new!**

- Another example: hh➜bbττ high mass analysis
  - dedicated high-$p_T$ h➜ττ reconstruction
  - spin-0, -1 (Vh➜qqττ), and -2 interpretation
- Also have analyses covering hh➜bbγγ (HIG-17-008), bblνlν (HIG-17-006) and low mass bbττ (HIG-16-002)

**spin-0 and -2 interpretation**

# More on di-Higgs production

**spin-0 interpretation**



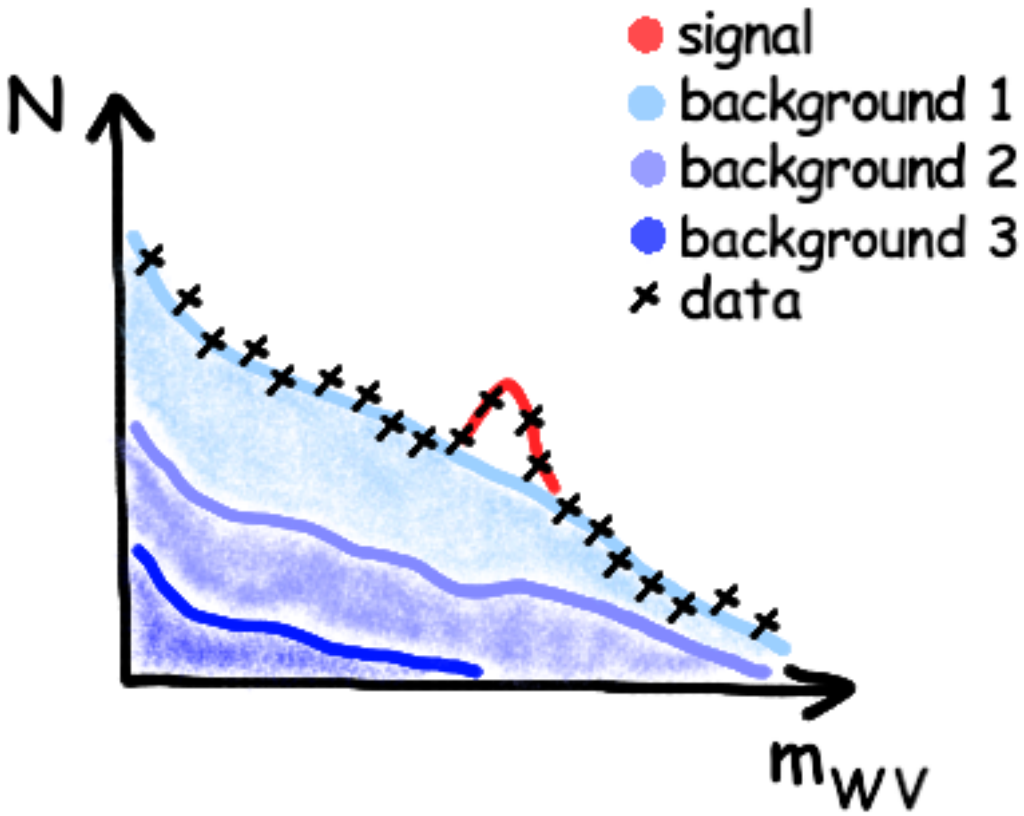
| | |
|---|---|
| bblνlν (HIG-17-006) | bbbb (HIG-17-009) |
| bbττ (HIG-17-002, submitted to Phys. Lett. B) | bbbb (B2G-16-026, submitted to Phys. Lett. B) |
| bbττ (B2G-17-006, submitted to J. High Energy Phys.) | bbγγ (HIG-17-008) |

**spin-2 interpretation in backup**

# (Semi-)leptonic VV final states

# (Semi-)leptonic final states

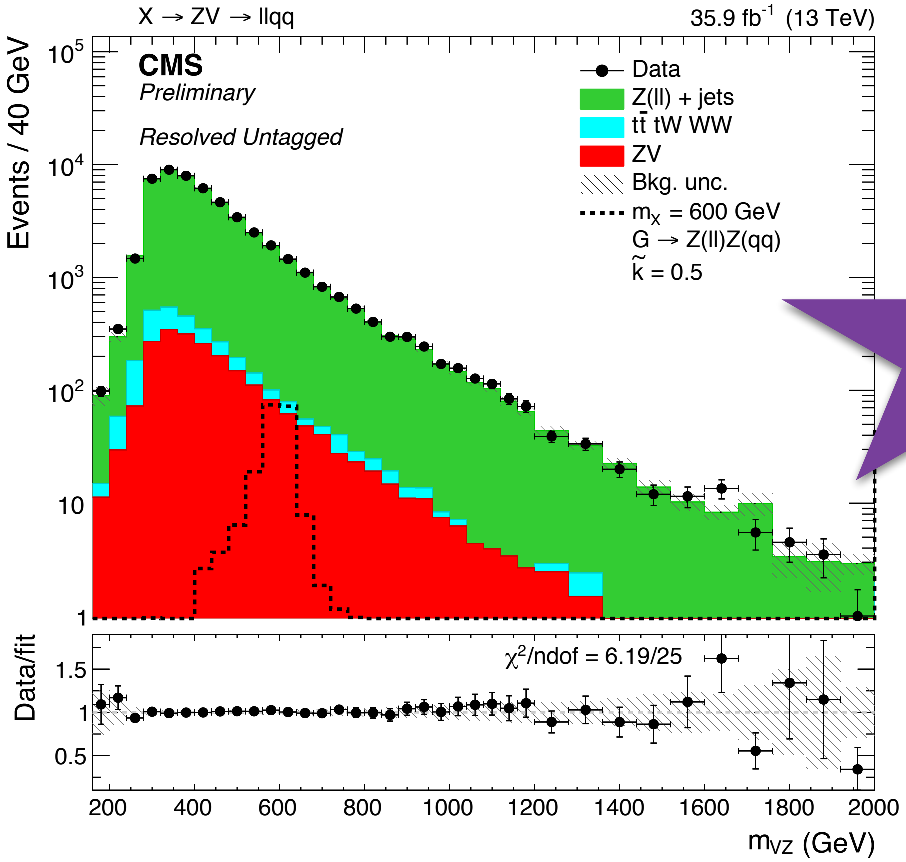
- Requiring one or more lepton(s) in the final state reduces background significantly
  - covering large number of final states (lν2q, 4l, 2l2ν, 2q2l, 2q2ν) and large mass range
- Lower trigger thresholds
- Simulation-assisted background estimation performed
  - distinguish different background processes
  - ratios for sideband to signal region extrapolation
  - or directly using simulation
- All analyses combined cover mass range from 130 to 4500 GeV



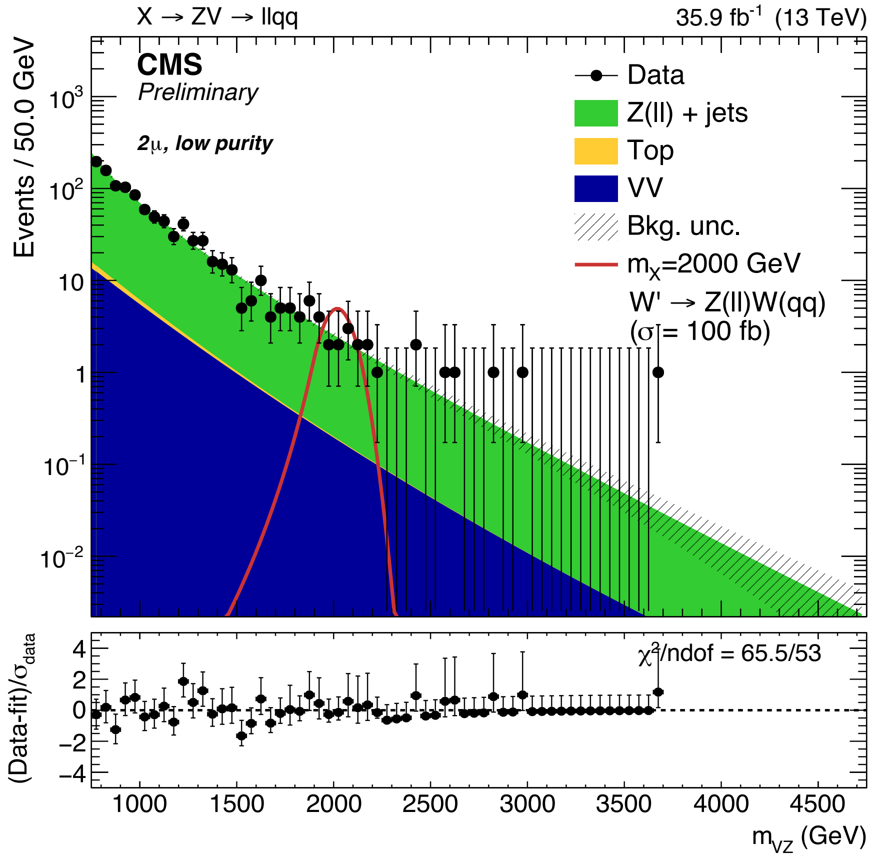
# ZV/ZZ (Z➔νν/ll) analyses

B2G-17-013

- Several different analyses:
  - high mass spin-2: 2l2ν (B2G-16-023)
  - high mass spin-1/-2: 2q2ν (B2G-17-005)
  - intermediate to high mass spin-1/-2: 2l2q (B2G-17-013)
  - low to high mass spin-0: 4l, 2l2ν, 2l2q (HIG-17-012)
- Recently released: B2G-17-013
  - 2l2q final state
  - dedicated low and high mass analyses
  - analysis extends from 400 GeV up to 4.5 TeV



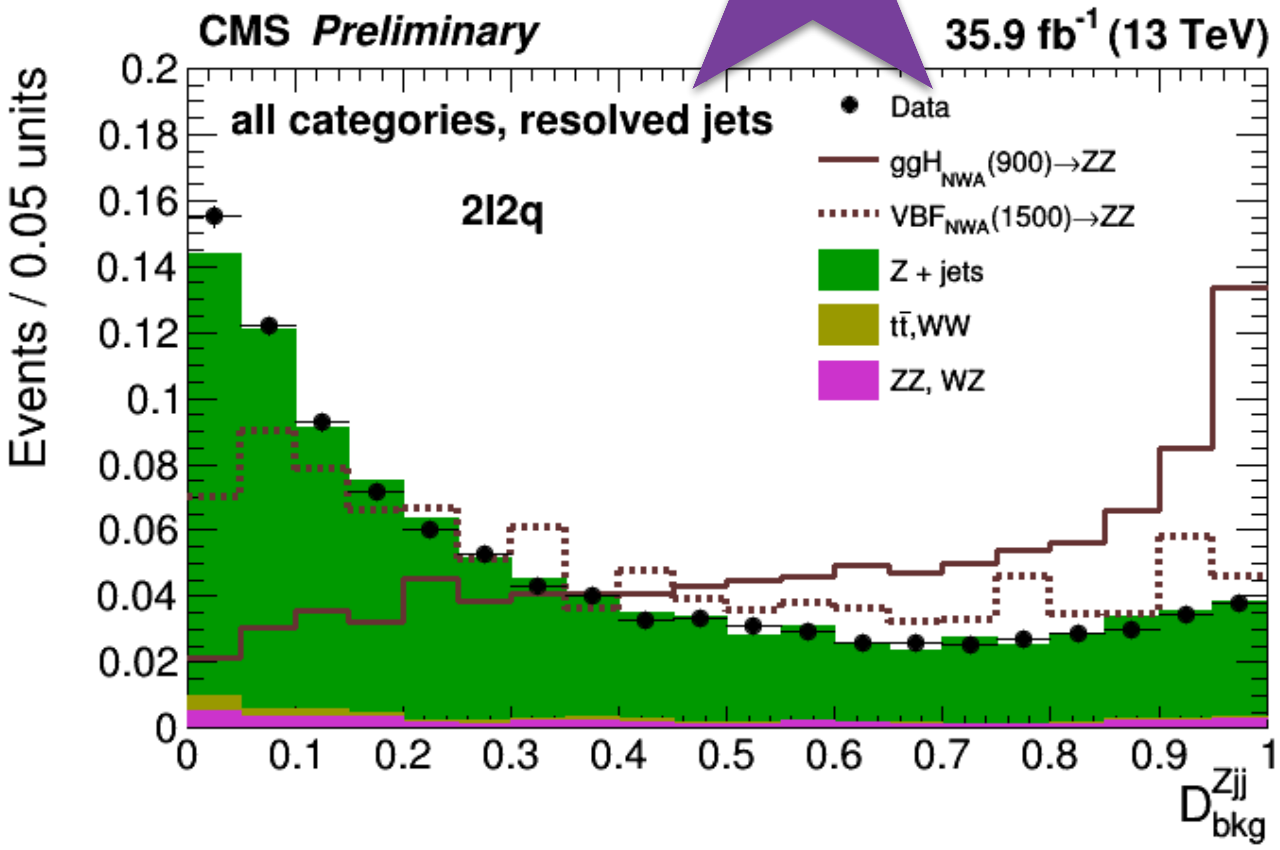
new!

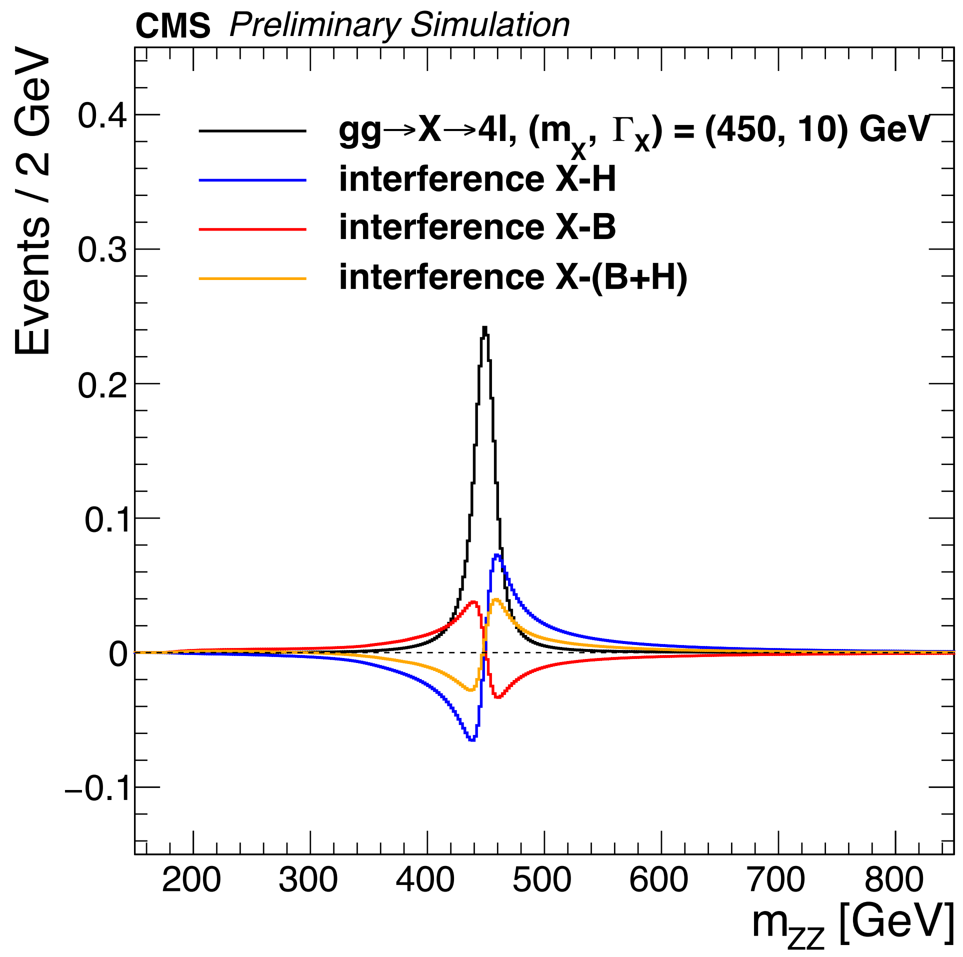


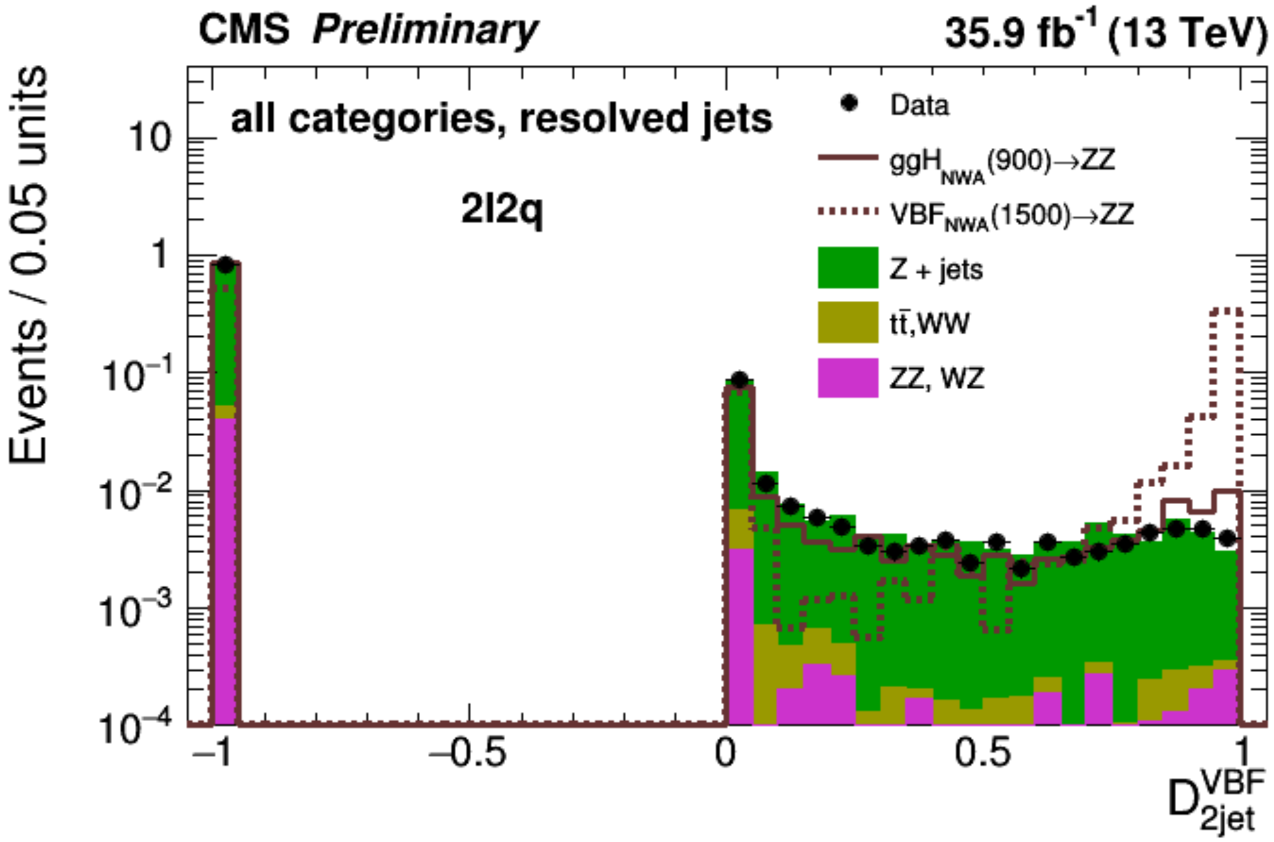
# ZV/ZZ (Z➔νν/ll) analyses

new!

- Recently released: HIG-17-012
- Combination of 4l, 2l2ν, and 2l2q
- MELA (Matrix Element Likelihood Analysis) discriminators for 4l & 2l2q (for spin 0 & spin 2)
- Considers S+B interference



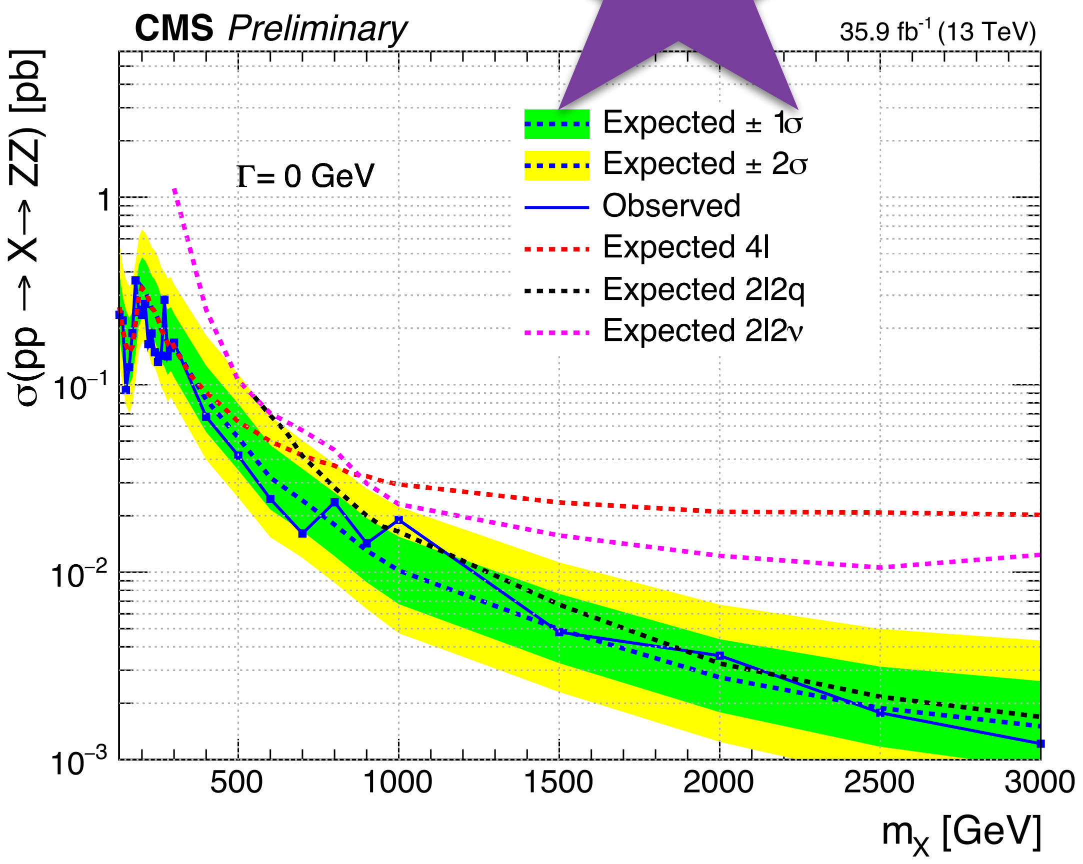


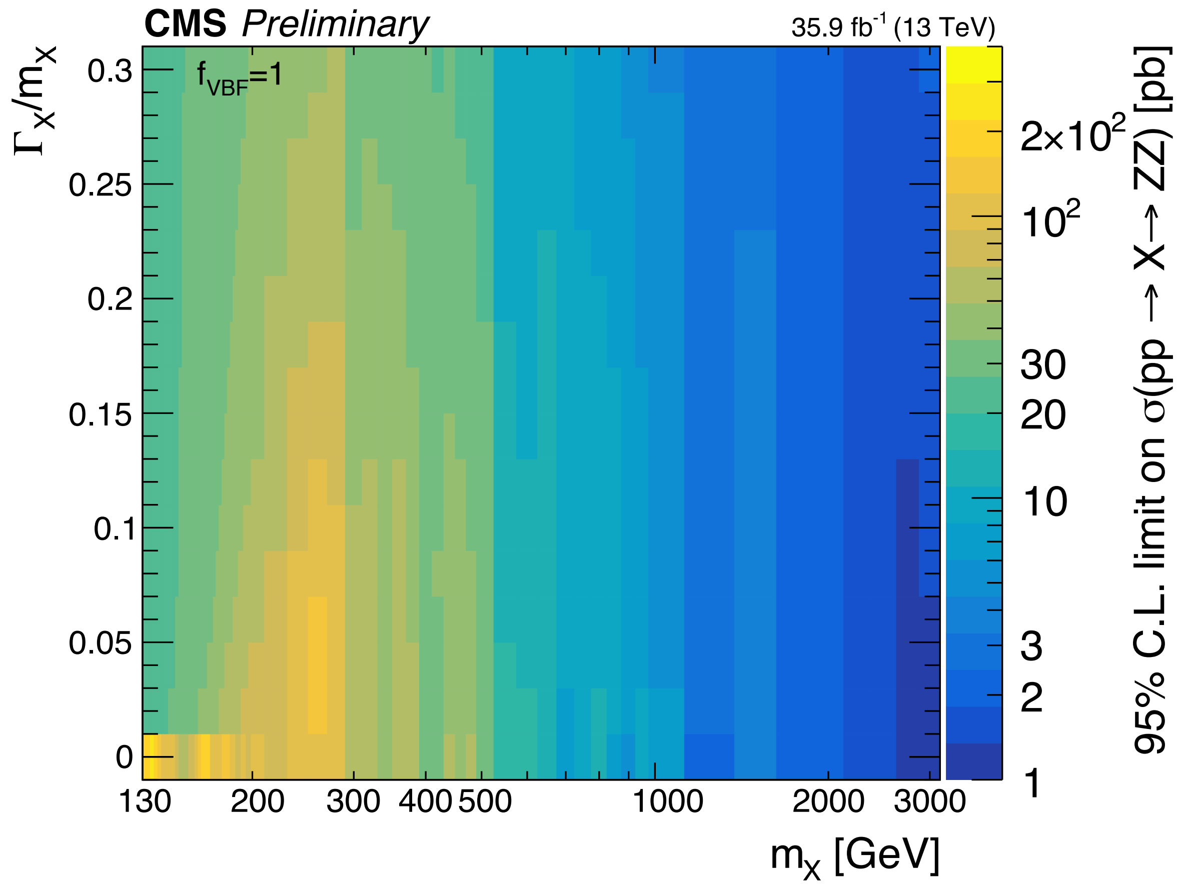


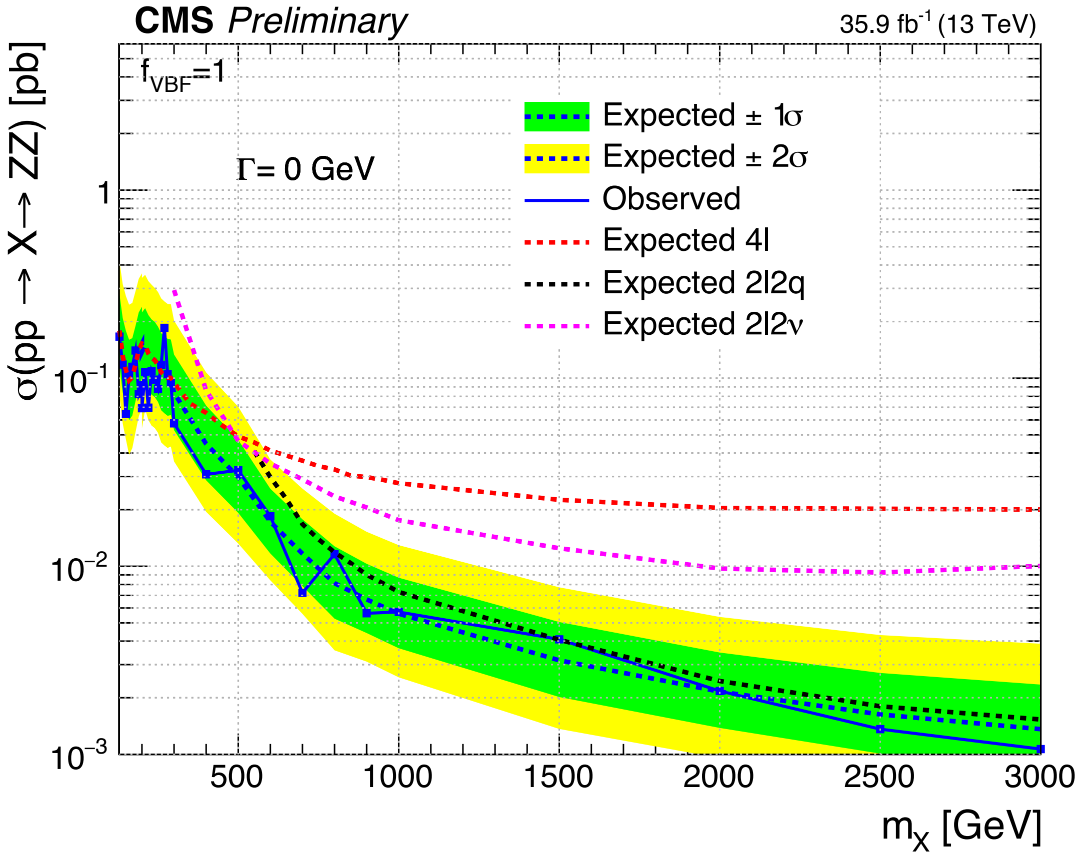
# ZV/ZZ (Z➔vv/ll) analyses

new!

- Considers ggH and VBF production
- Model-independent limits from 130 to 3000 GeV
  - up to 30% resonance width



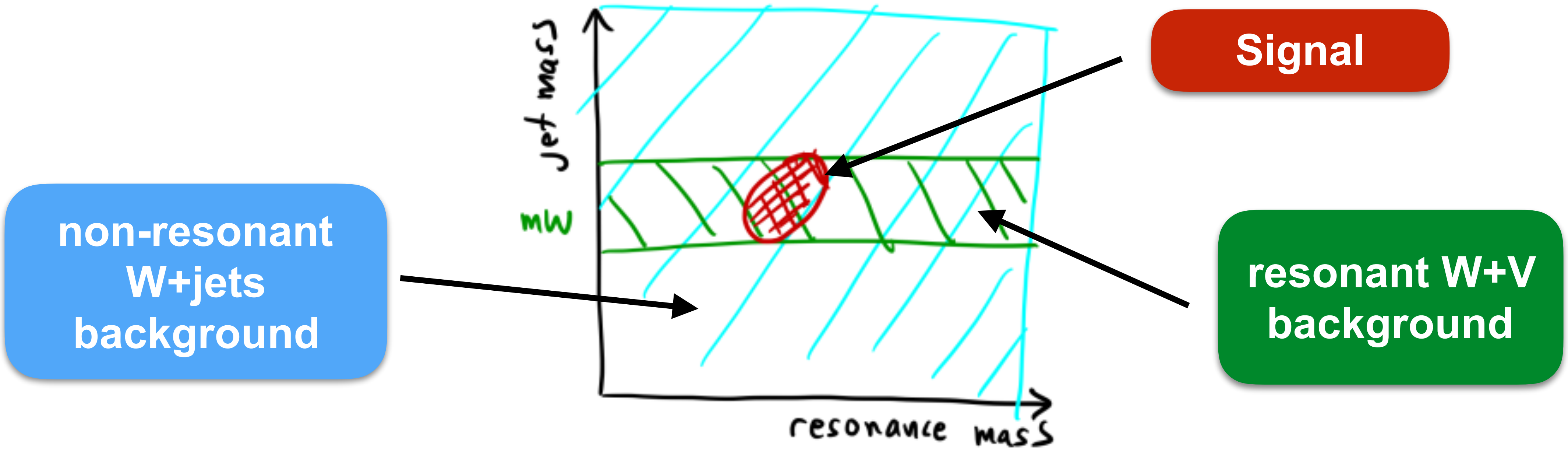




# WV (W➜lν) analysis

new!

- New analysis approach for WV➜lνqq analysis
  - do not use jet mass windows anymore
- Instead: 2D fit in ($m_{WV}$, $m_{jet}$) plane - use full V jet mass range: 30 < $m_{jet}$ < 210 GeV
  - make better use of correlations between $m_{WV}$ and $m_{jet}$
  - much more sideband statistics — use full line-shape of jet mass
  - become less dependent on simulation — learn from data
- 2D fit: distinguish between
  - non-resonant W+jets (W(lν)+jets, ttbar with non-W V jet)
  - resonant W+V (ttbar, diboson) background processes



**Signal**

**non-resonant W+jets background**

**resonant W+V background**

# WV analysis parametrisation

new!

- Using conditional probabilities for signal and each background component starting from simulation
  - taking into account scale and resolution
- Signal peaks in both $m_{WV}$ and $m_{jet}$:

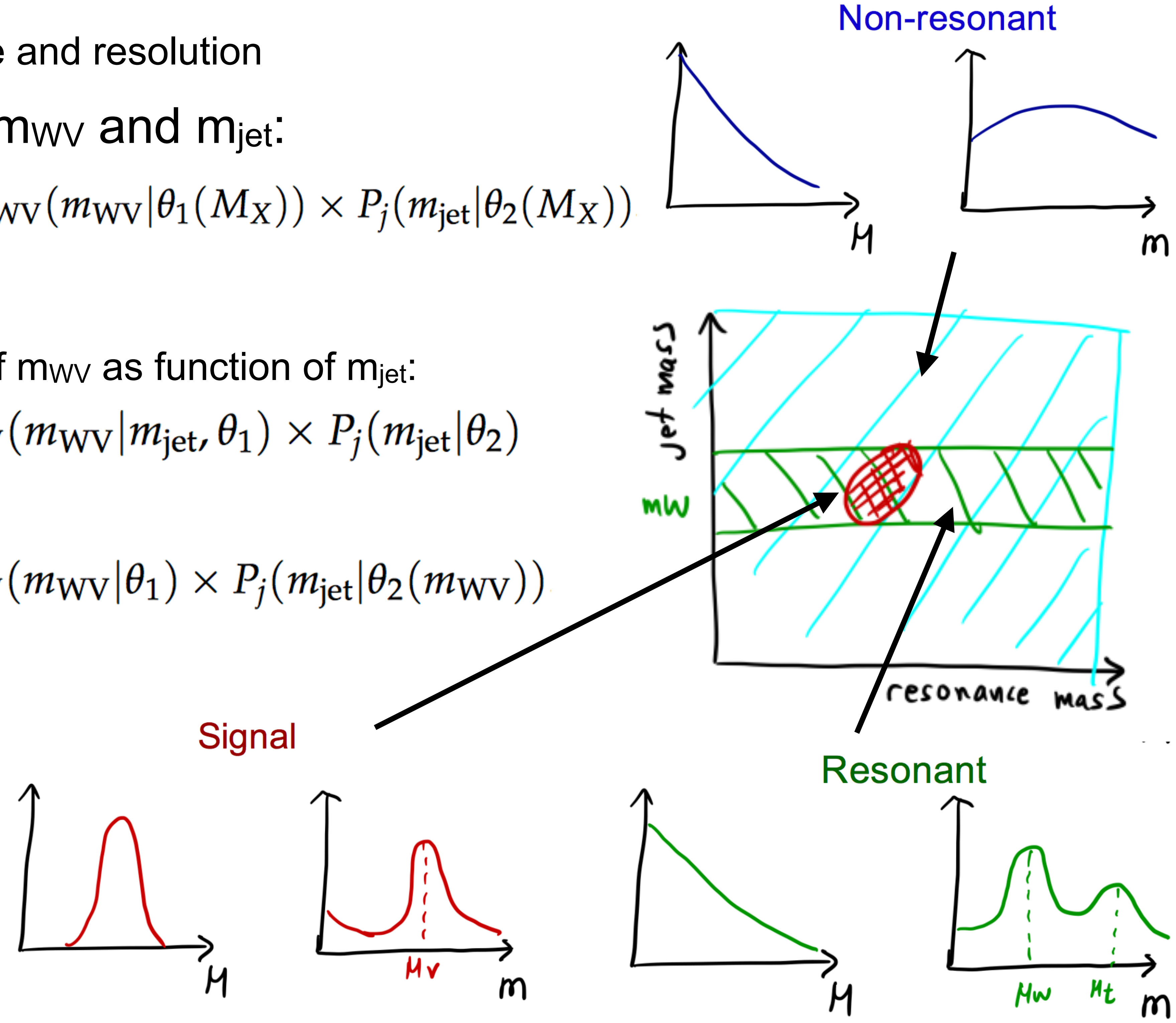
$$P_{sig}(m_{WV}, m_{jet}|\theta(M_X)) = P_{WV}(m_{WV}|\theta_1(M_X)) \times P_j(m_{jet}|\theta_2(M_X))$$

- W+jets background:
  - Conditional probability of $m_{WV}$ as function of $m_{jet}$:

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

- W+V background:

$$P_{W+V}(m_{WV}, m_{jet}|\theta) = P_{WV}(m_{WV}|\theta_1) \times P_j(m_{jet}|\theta_2(m_{WV}))$$



Non-resonant

Signal

Resonant

# Post-fit projections

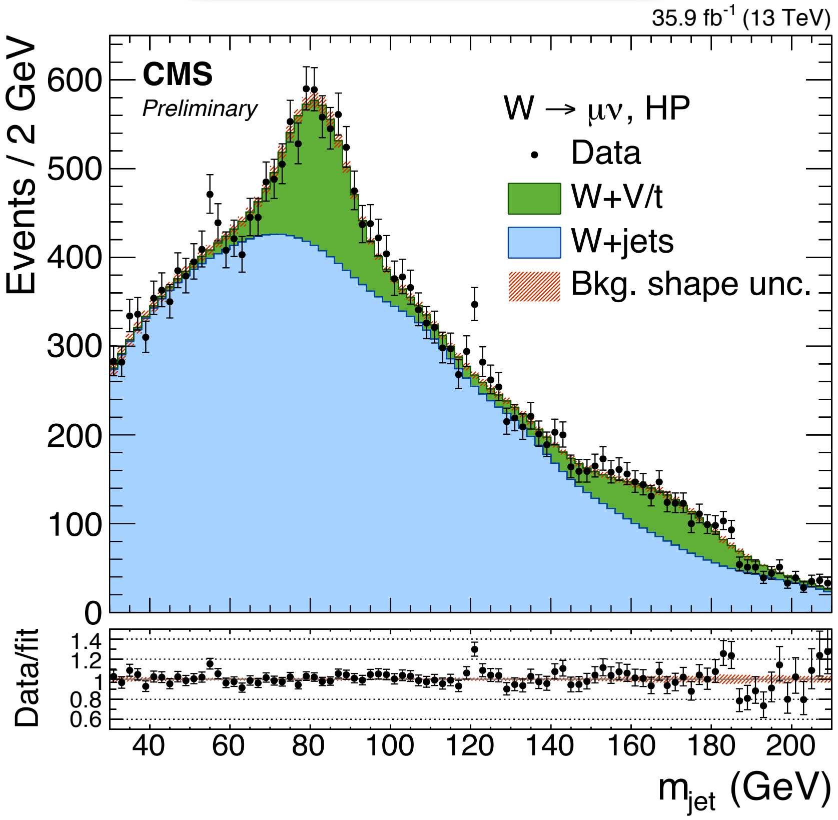
B2G-16-029

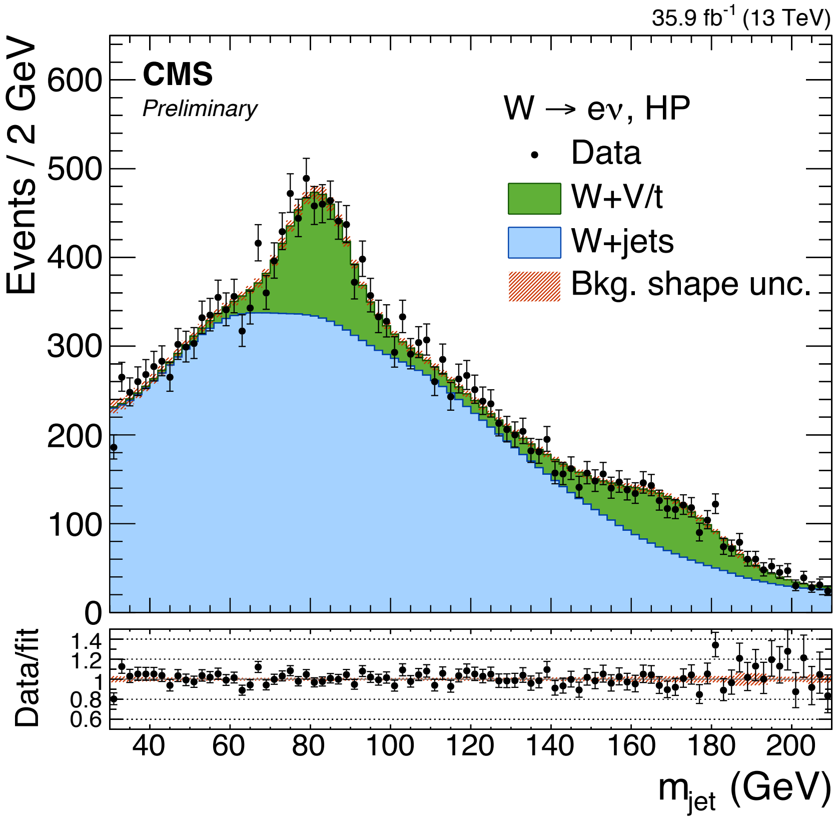
new!

**muon**

**electron**

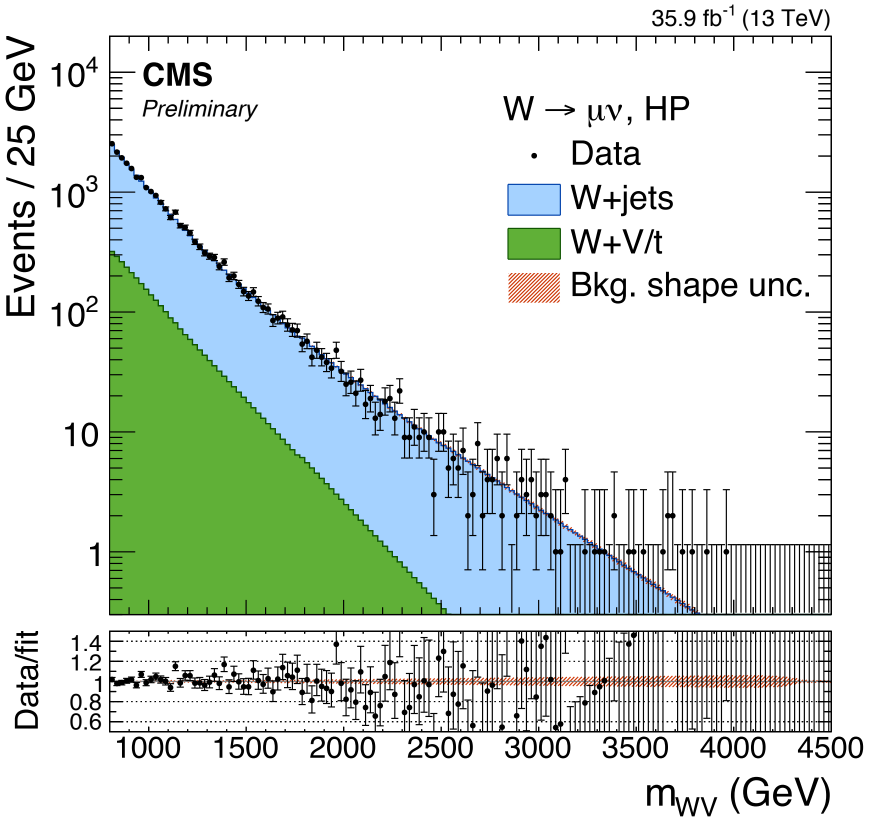
**$m_{jet}$**

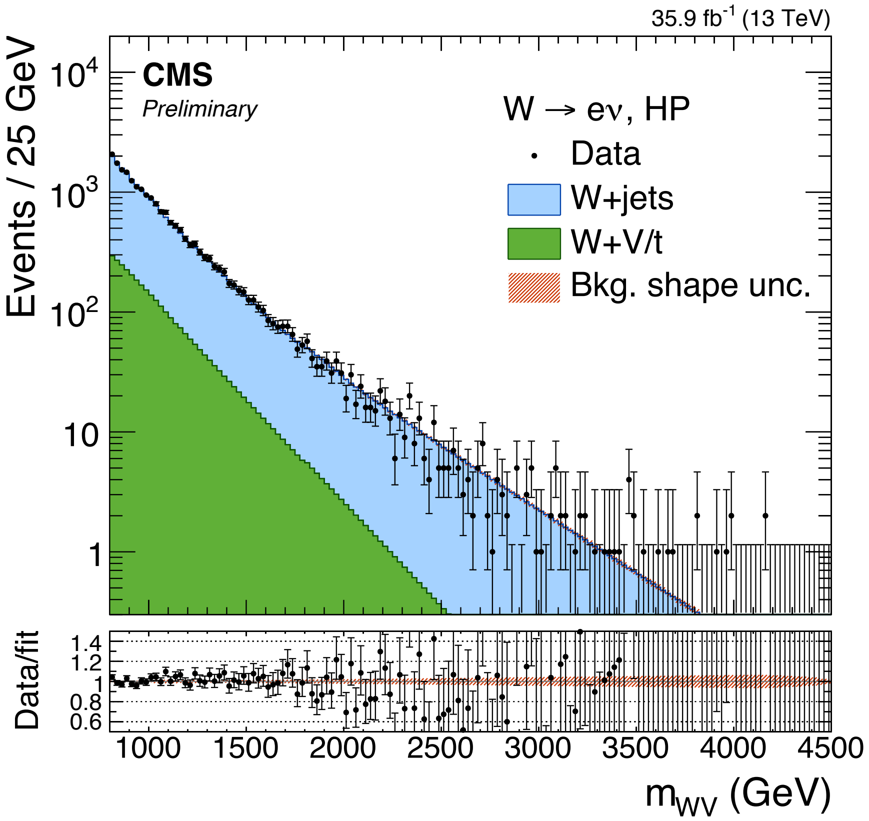




**Nicely modelling both W and top peaks!**

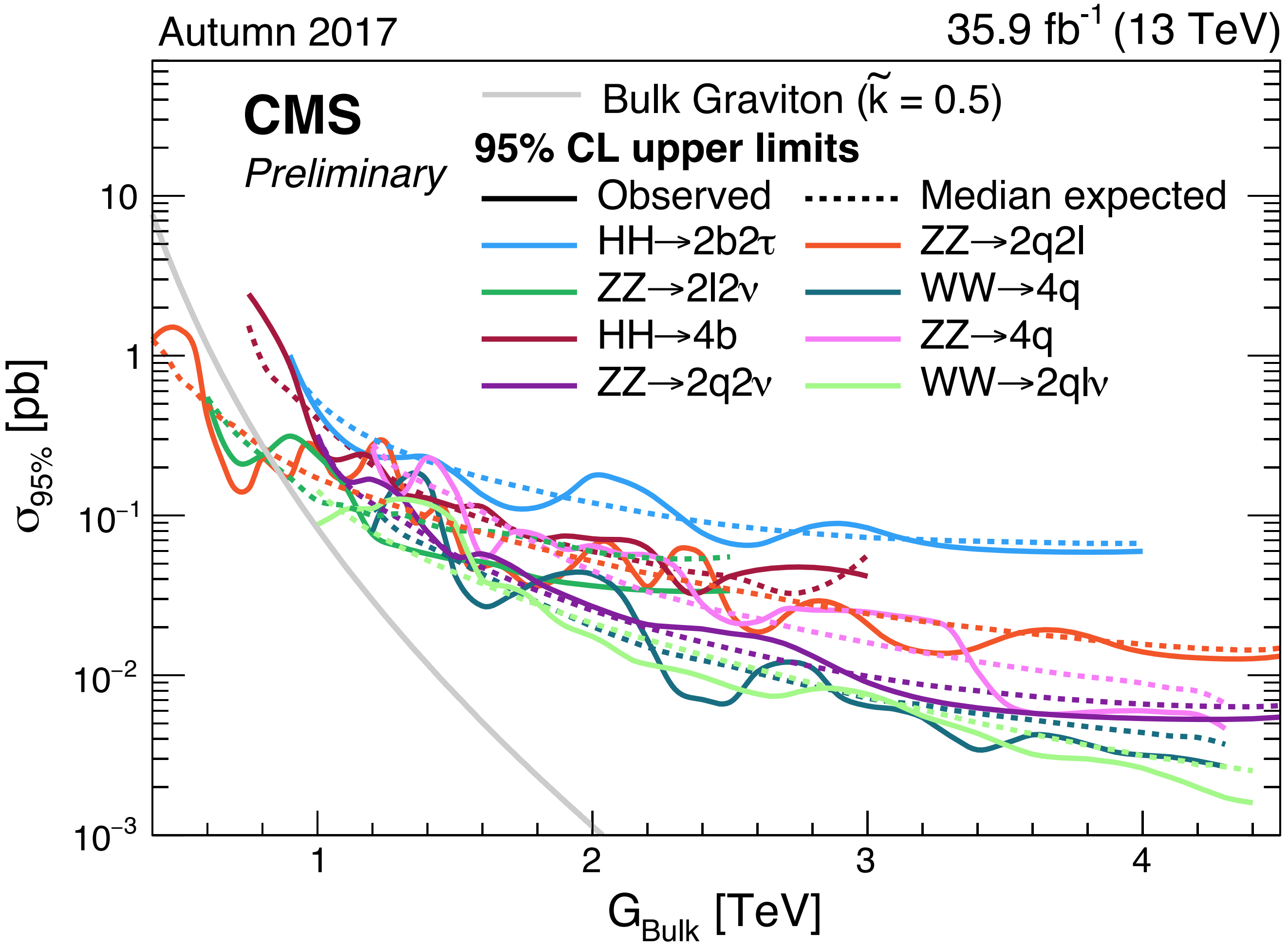
**$m_{WV}$**



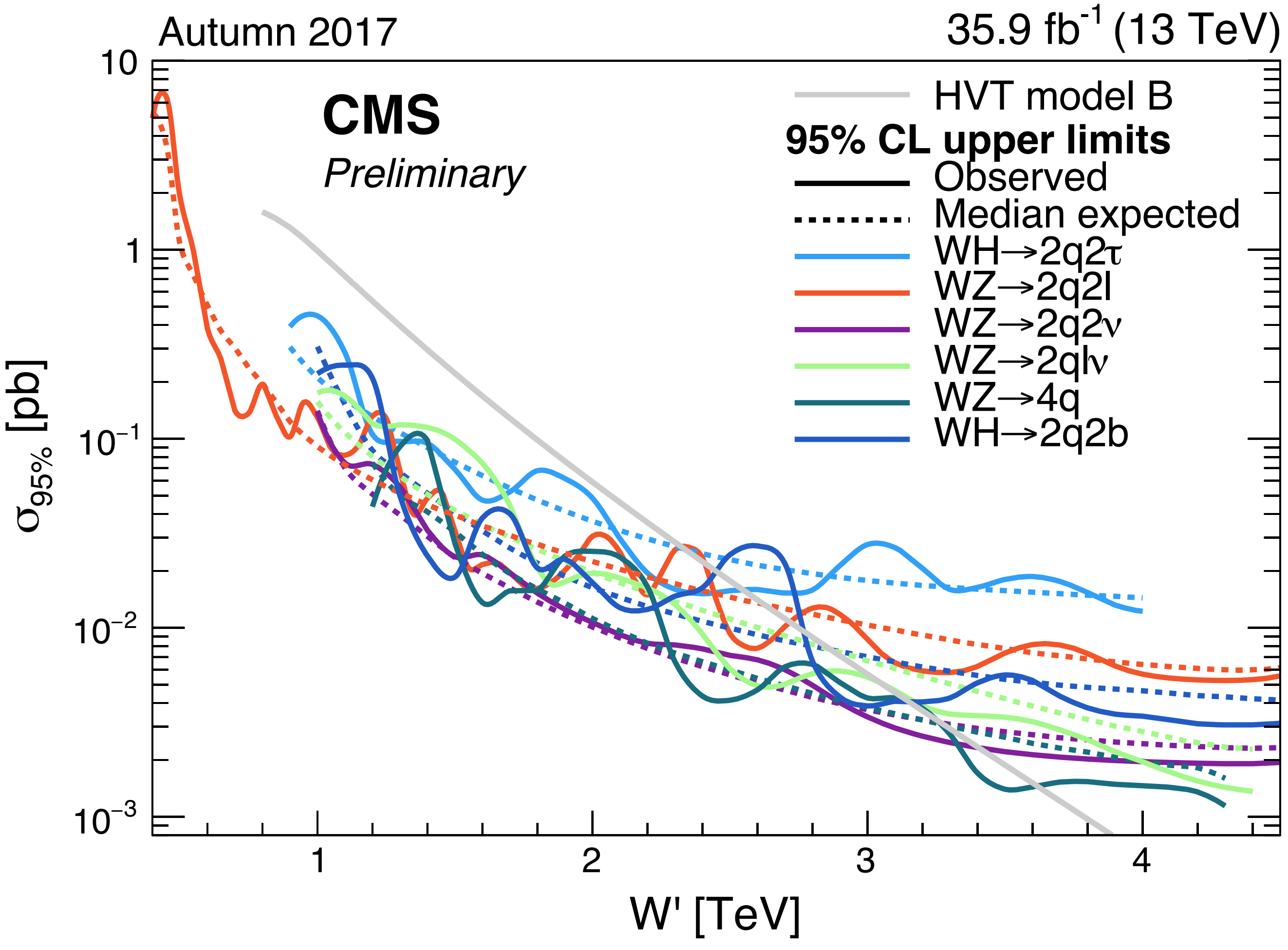


# High mass diboson resonances summary

**bulk graviton (spin-2)**



**heavy vector triplet (W', spin-1)**



HH➔2b2τ/WH➔2q2τ (B2G-17-006)

ZZ/WZ➔2q2l (B2G-17-013)

ZZ➔2l2ν (B2G-16-023, submitted to J. High Energy Phys.)

WW/WZ/ZZ➔4q (B2G-17-001, submitted to Phys. Rev. D)

WW/WZ➔2qlν (B2G-16-029)

ZZ➔2q2ν (B2G-17-005)

HH➔4b (B2G-16-026, submitted to Phys. Lett. B)

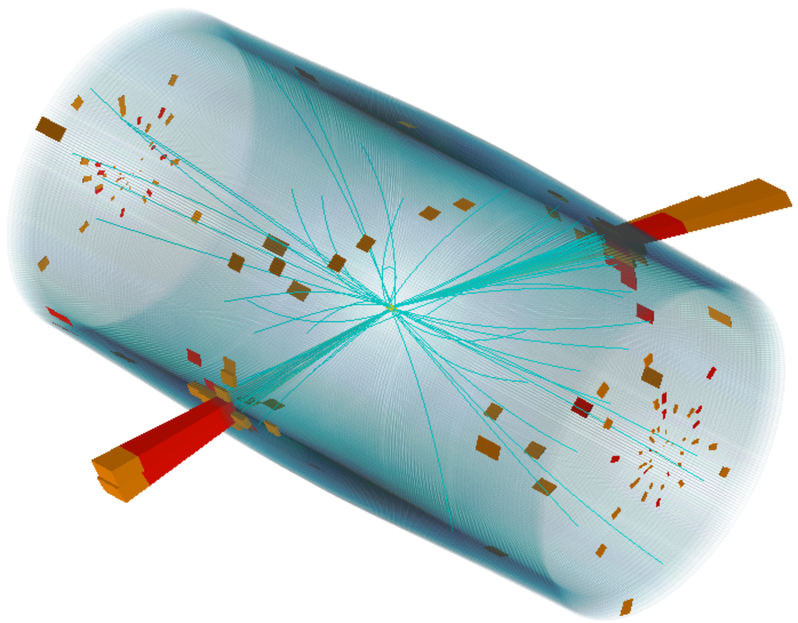
WH➔2q2b (B2G-17-002, Eur. Phys. J. C 77 (2017) 636)

# Challenges ahead & conclusions

Challenges:

- Run-2 data set will yield higher statistics, but no increase in centre-of-mass energy
- Imperative to further improve methodology
  - multi-dimensional fits make best use of statistics
- Further work needed on understanding jet substructure
  - new boson tagging algorithms on the market
  - can also profit from machine learning

**Candidate ZZ event**
Dijet mass: 3.2 TeV

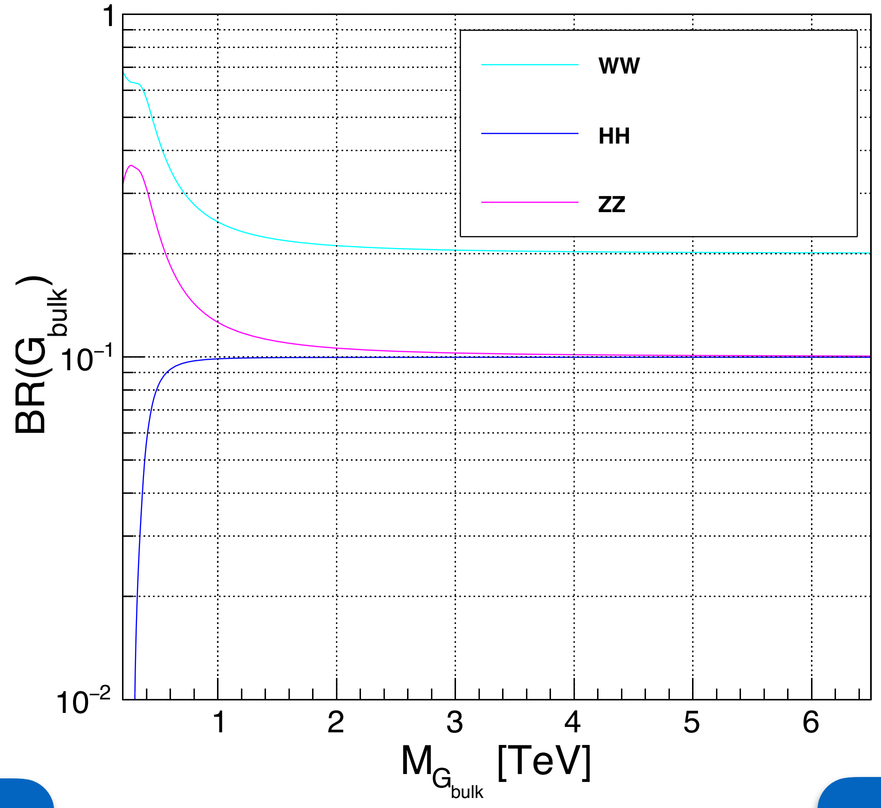


- CMS has an extensive diboson resonance search programme
- Several additions and improvements w.r.t. previous analyses
  - significantly higher sensitivity and extended mass range
- Well set up to make best use of full Run-2 data set


CERN

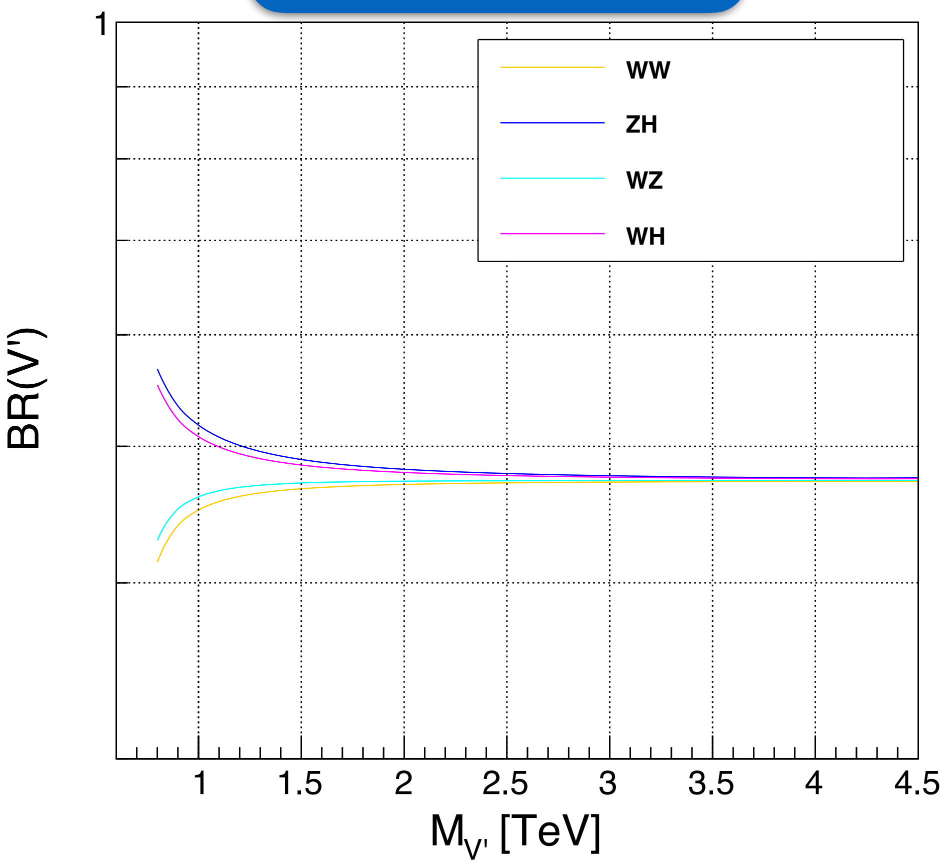
# Branching fractions

**bulk graviton**

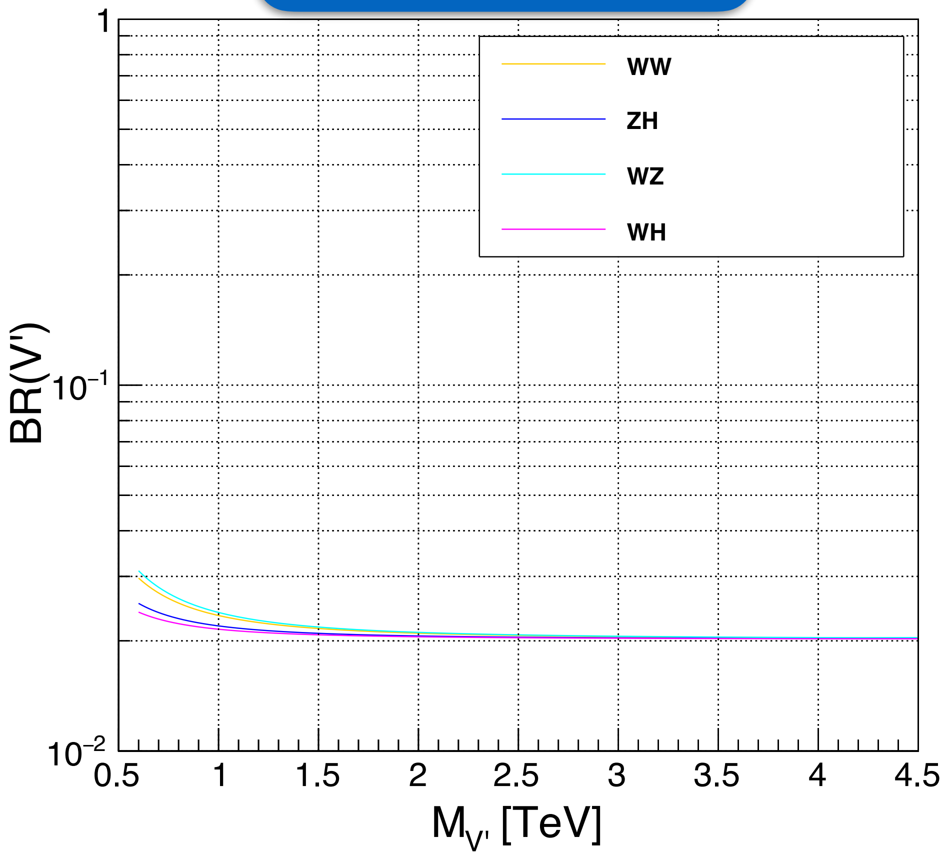


vector bosons are produced with a **longitudinal polarization** in more than **99%** of the cases, resulting in a **~24% higher acceptance per boson** than for models producing **transversally polarized** vector bosons

**HVT model B**

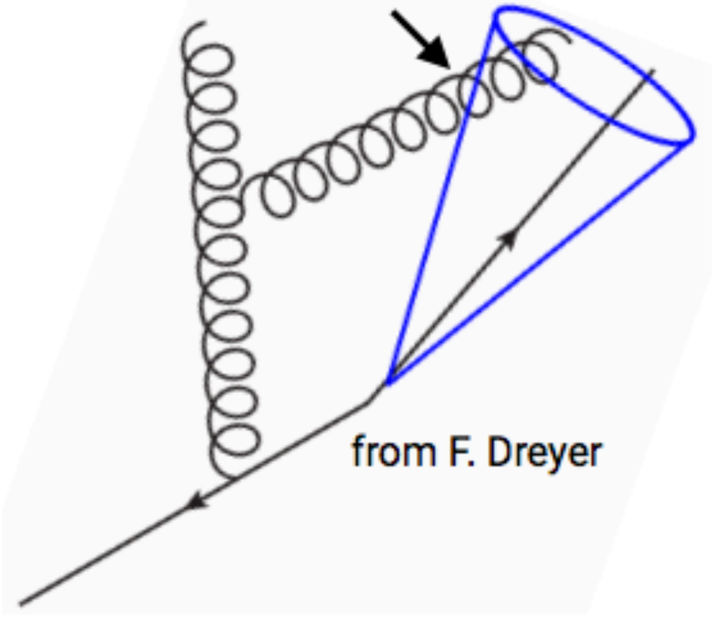


**HVT model A**

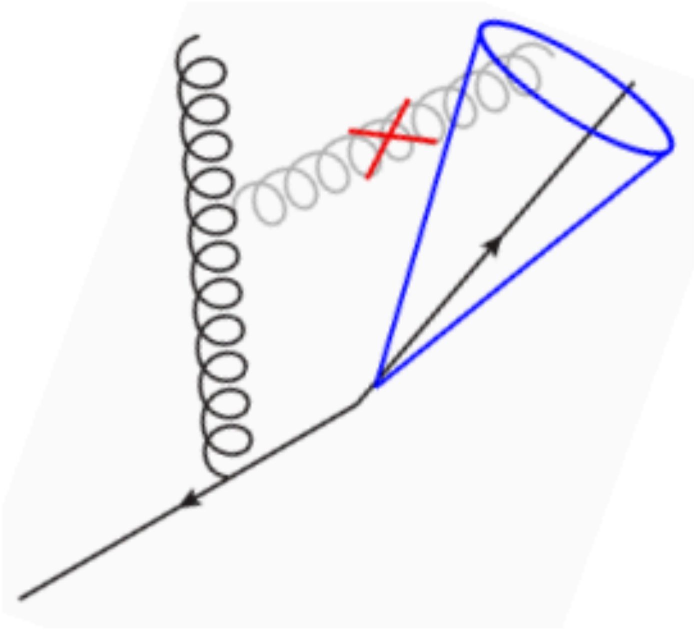


# Soft drop vs. pruning

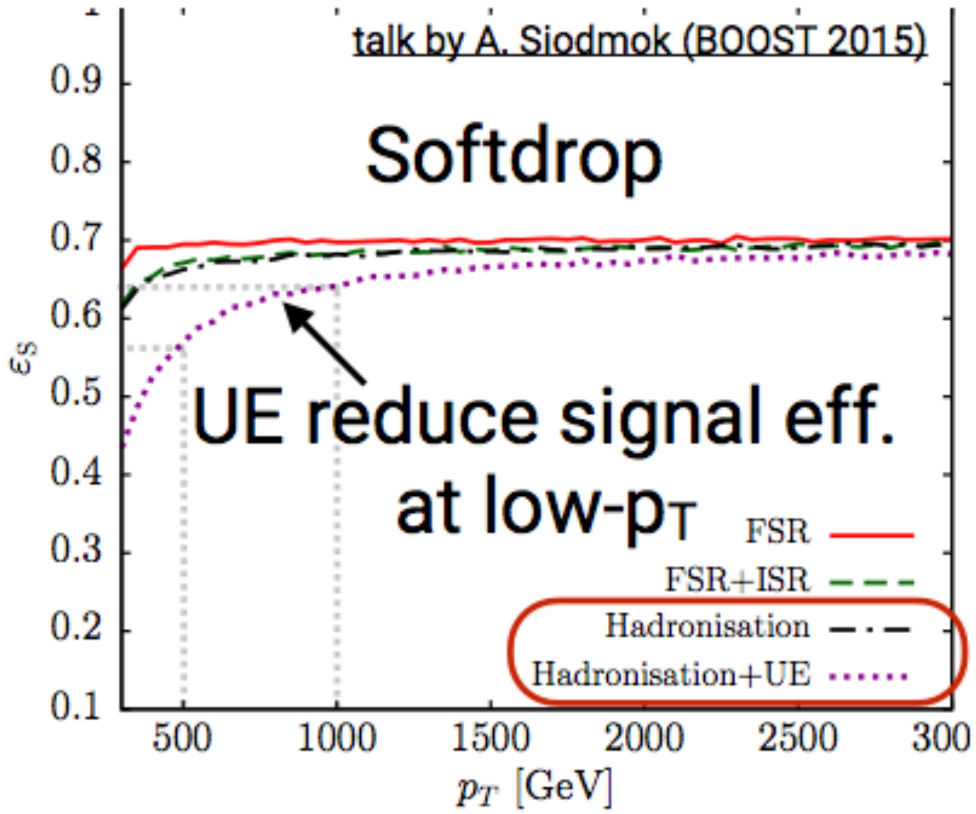
Pruning: Soft gluon radiating into jet not removed



Softdrop: All soft stuff removed. NGL free



- Pruning not "perturbatively robust"
  - not soft radiation free
  - nonglobal logarithmic terms (NGLs) in mass
- Soft drop:
  - removes all sensitivity to soft divergences
  - however, signal sensitive to underlying event (for multijets, pruning more sensitive to UE)
  - correct with PUPPI and mass corrections



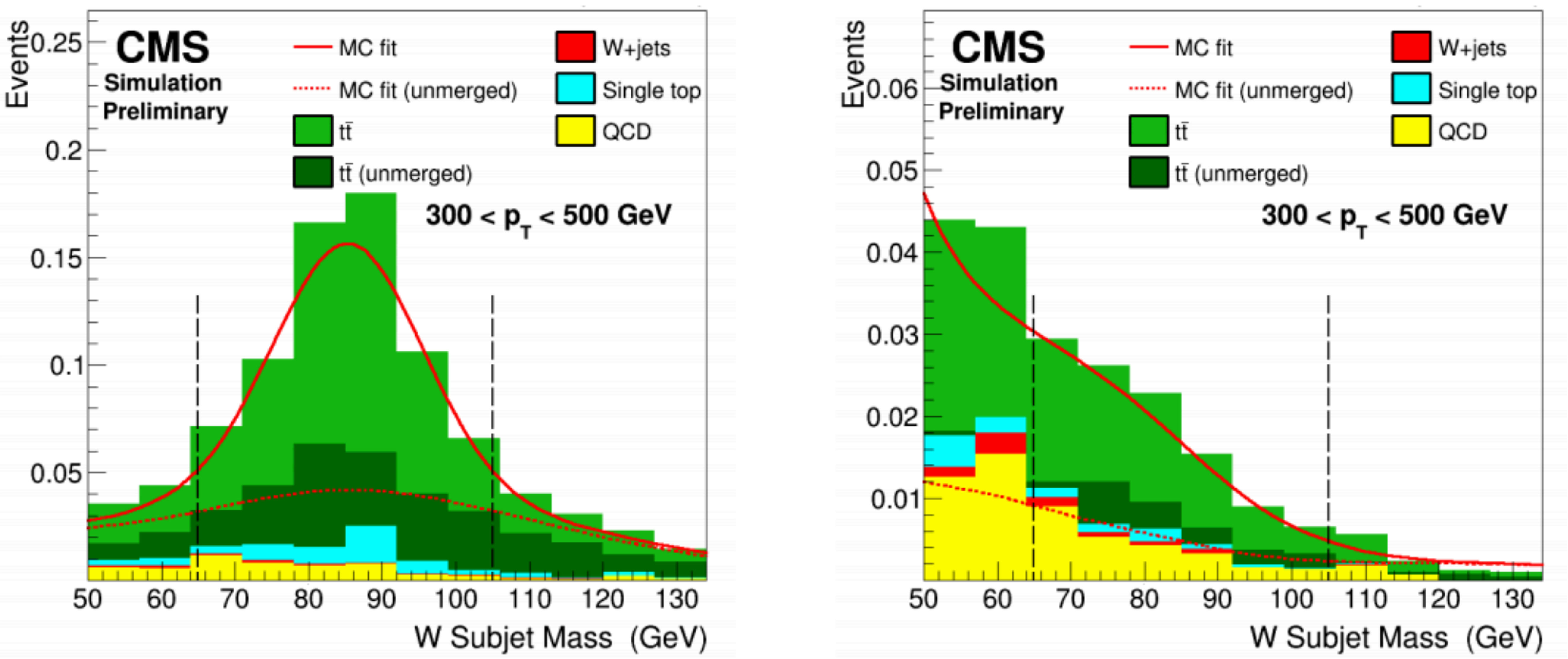
talk by A. Siodmok (BOOST 2015)

Softdrop

UE reduce signal eff. at low-$p_T$

# W-tagging at high $p_T$

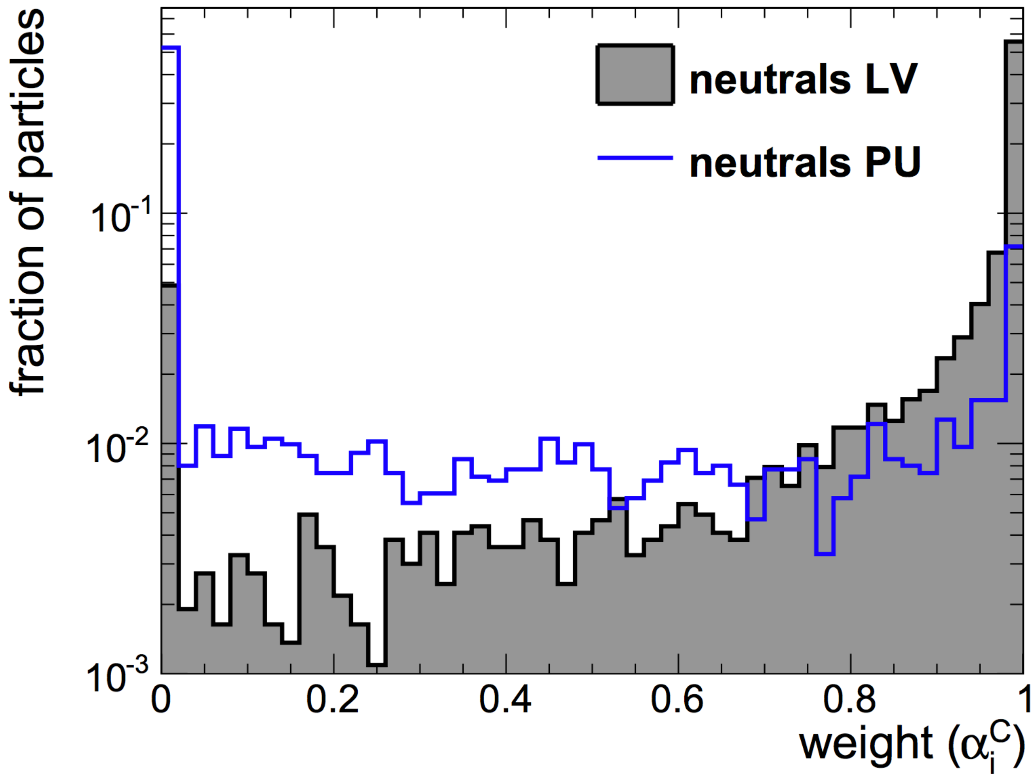
- Previously used simulation-based extrapolation of scale factors
- Now have data-based calibration available

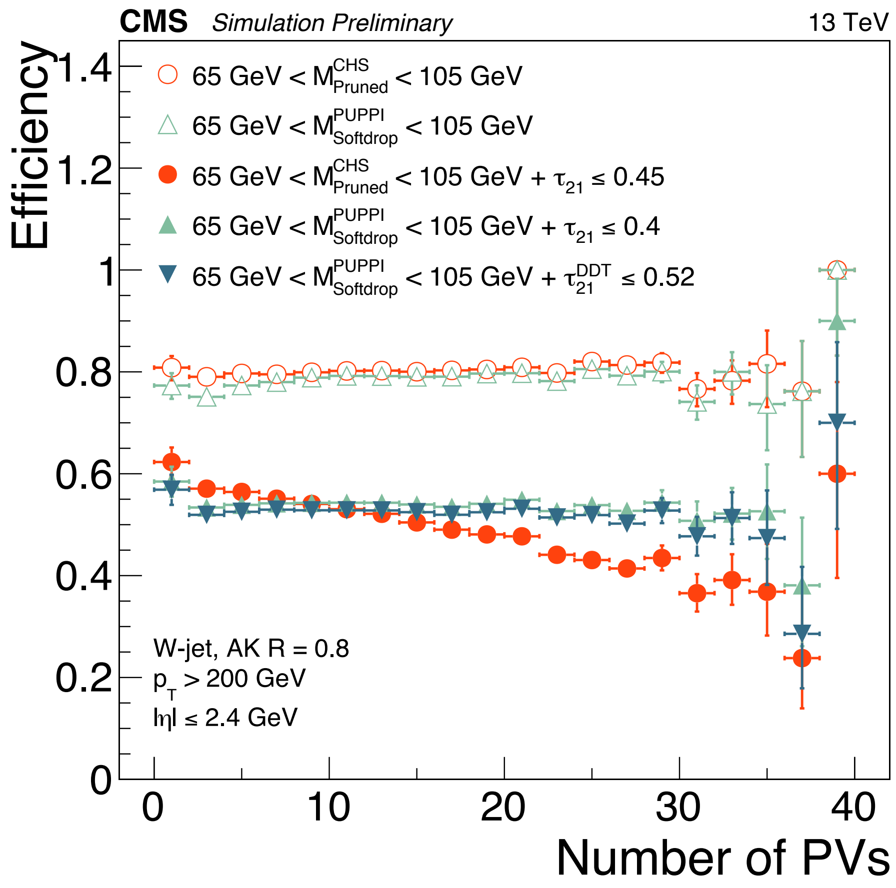


# PUPPI (PileUp Per Particle Identification)

JHEP 1410 (2014) 59
JME-16-003

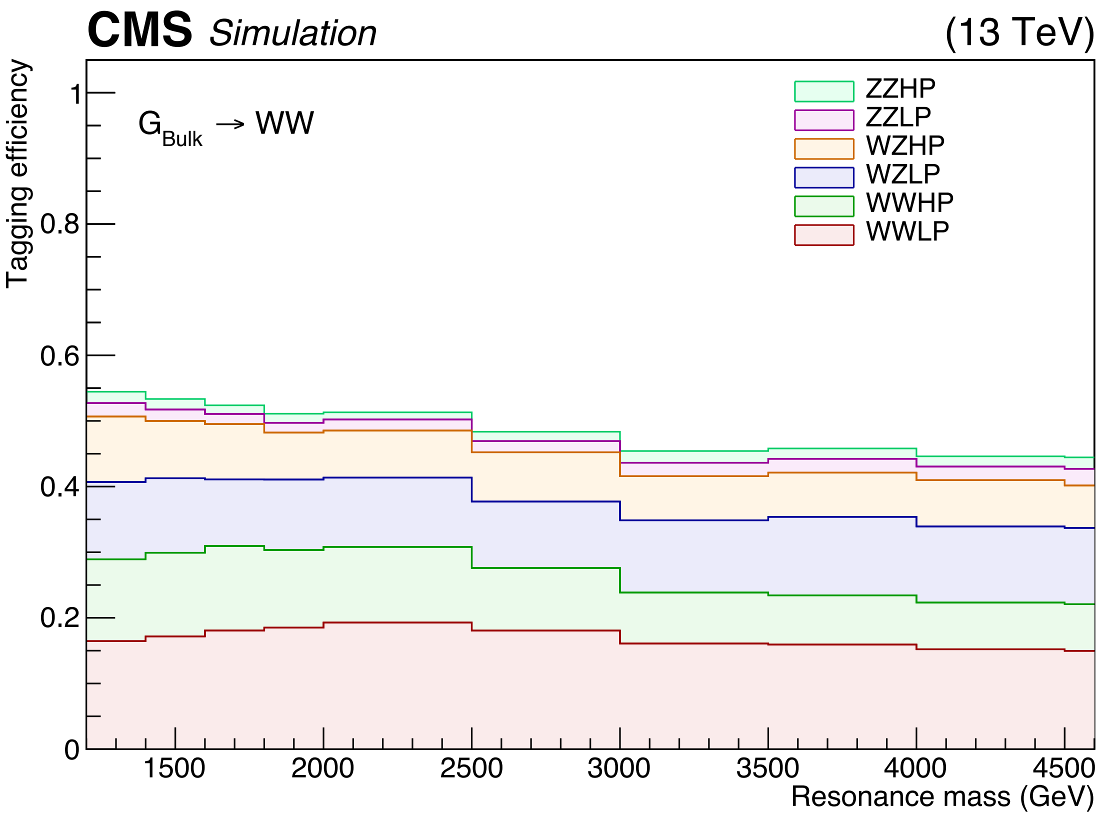
- Tracks can point to PU vertices, only keep charged tracks that come from the primary vertex
- Draw a cone around each neutral Particle Flow candidate
- Define a local metric, α, that differs between pileup (PU) and leading vertex (LV)
- For the neutrals, ask “how PU-like is α for this particle?”, compute a weight for how LV-like it is
- Reweight the four-vector of the particle by this weight

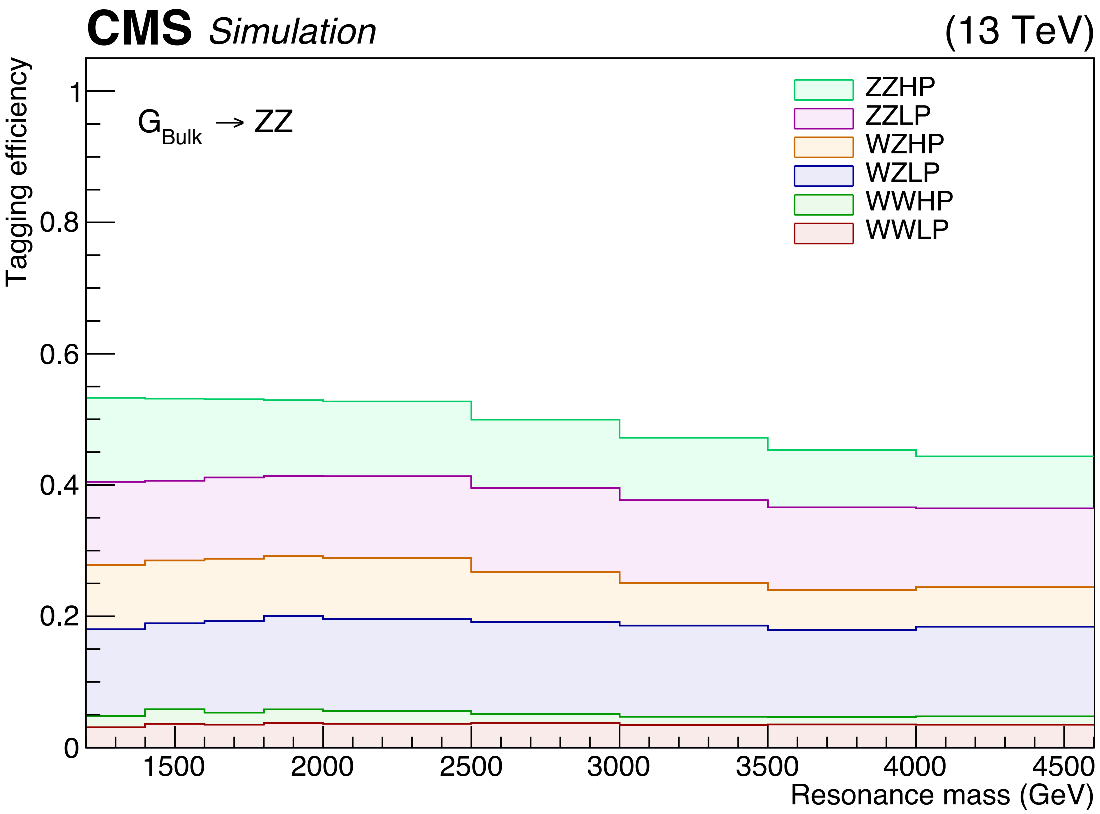


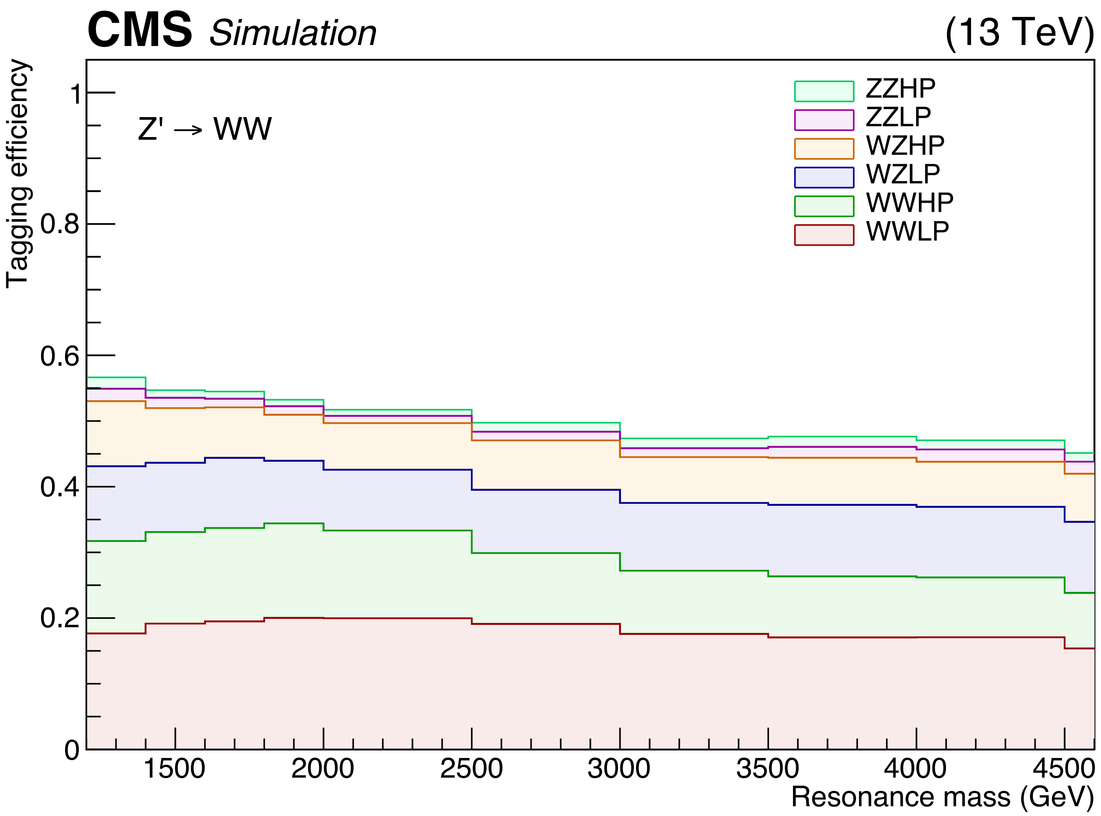


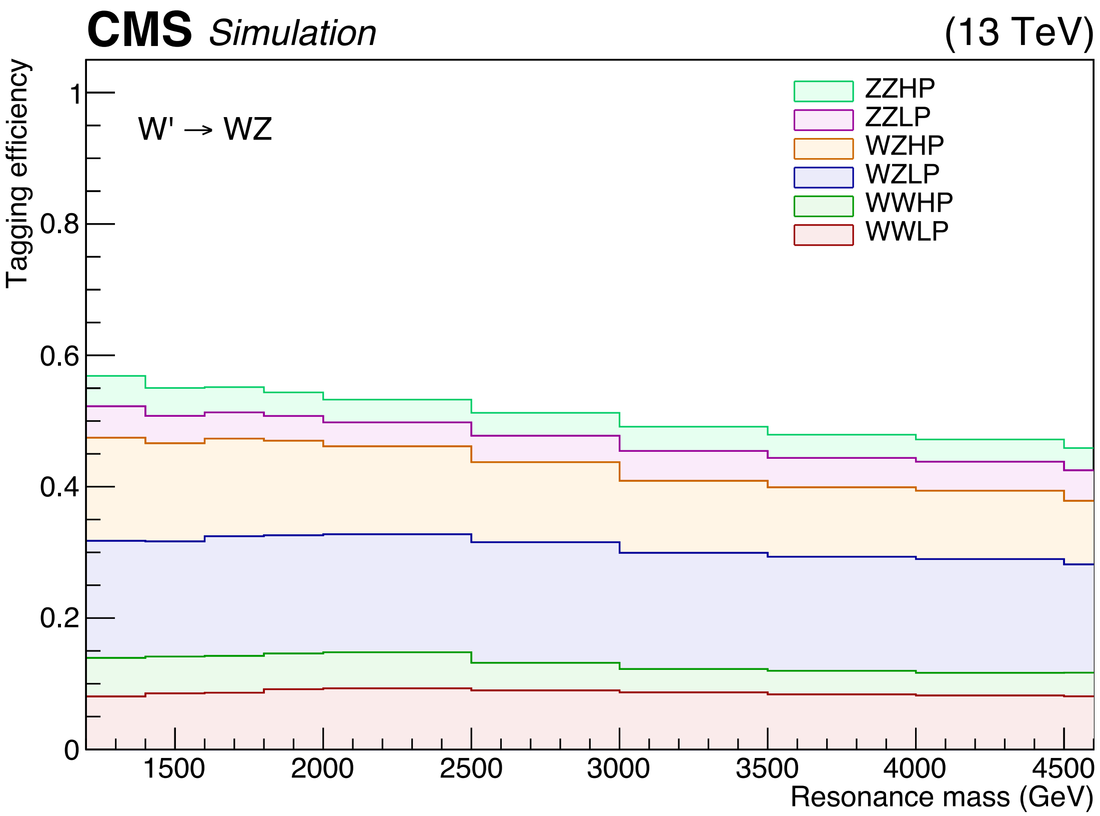
# V-tagging efficiencies

B2G-17-001 (submitted to Phys. Rev. D)

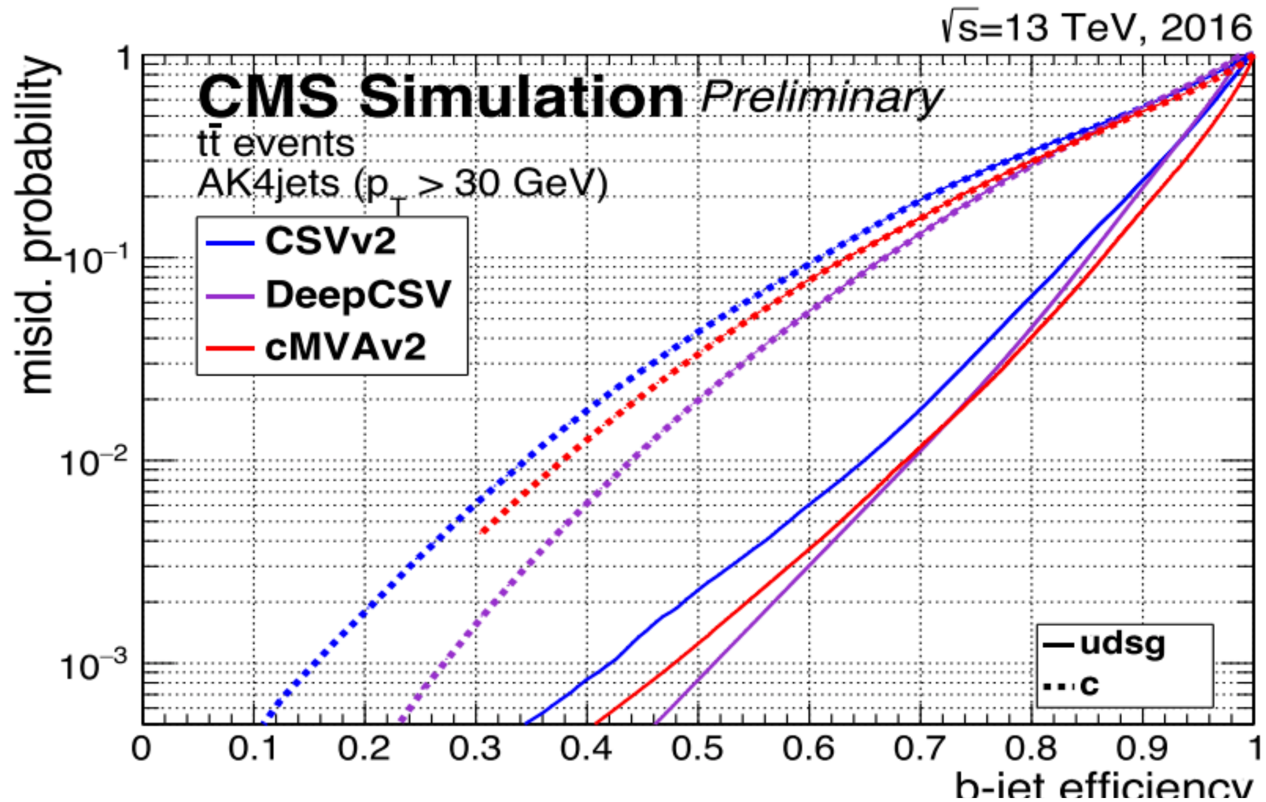


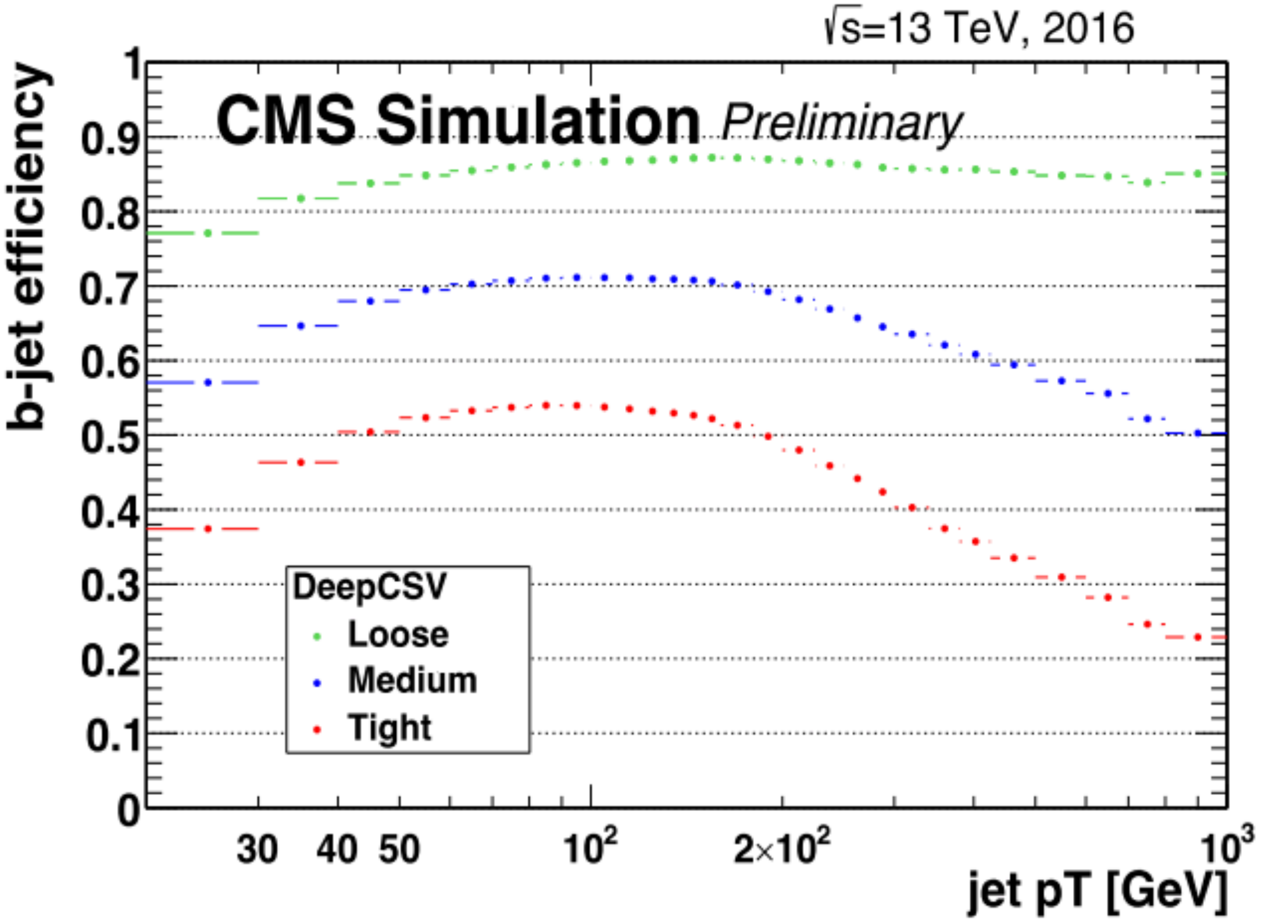


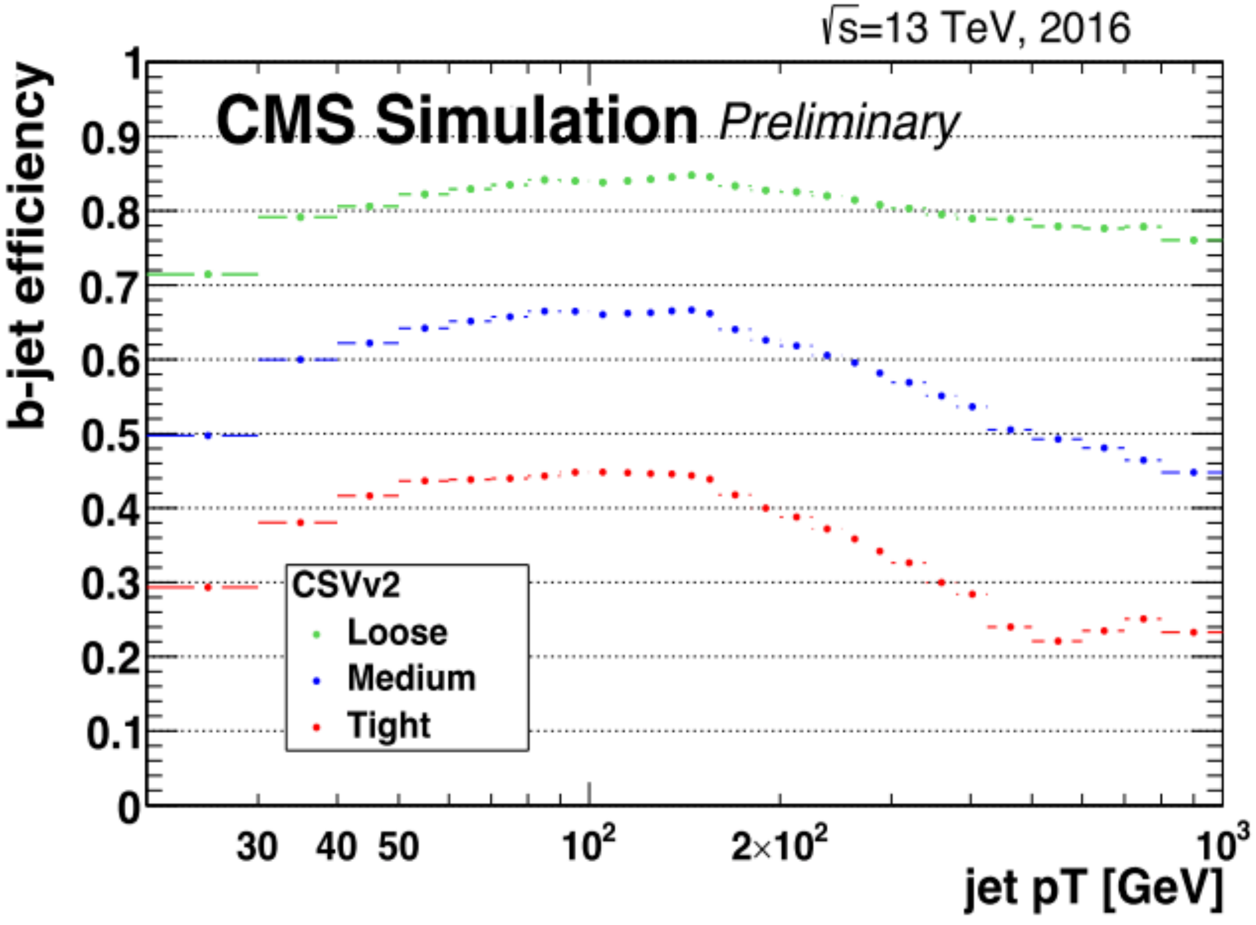




# DeepCSV performance



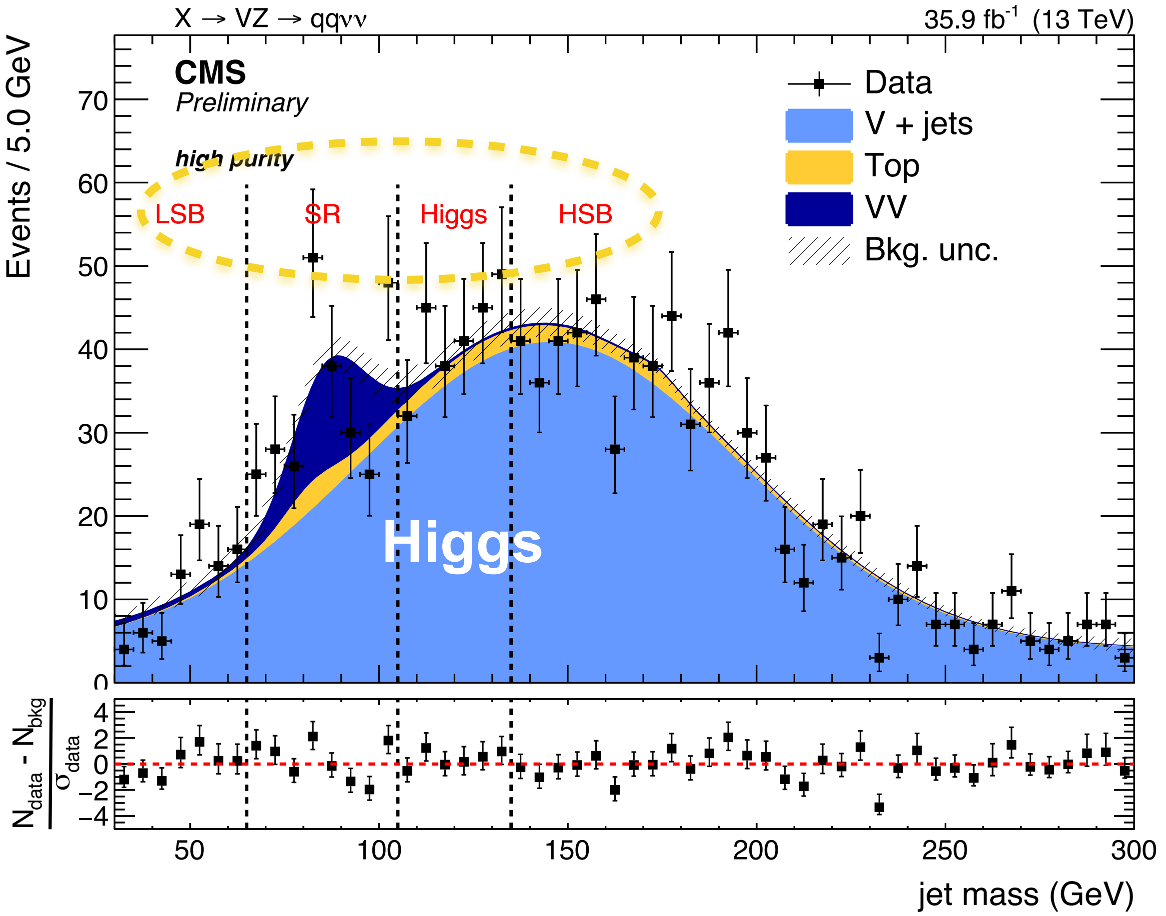


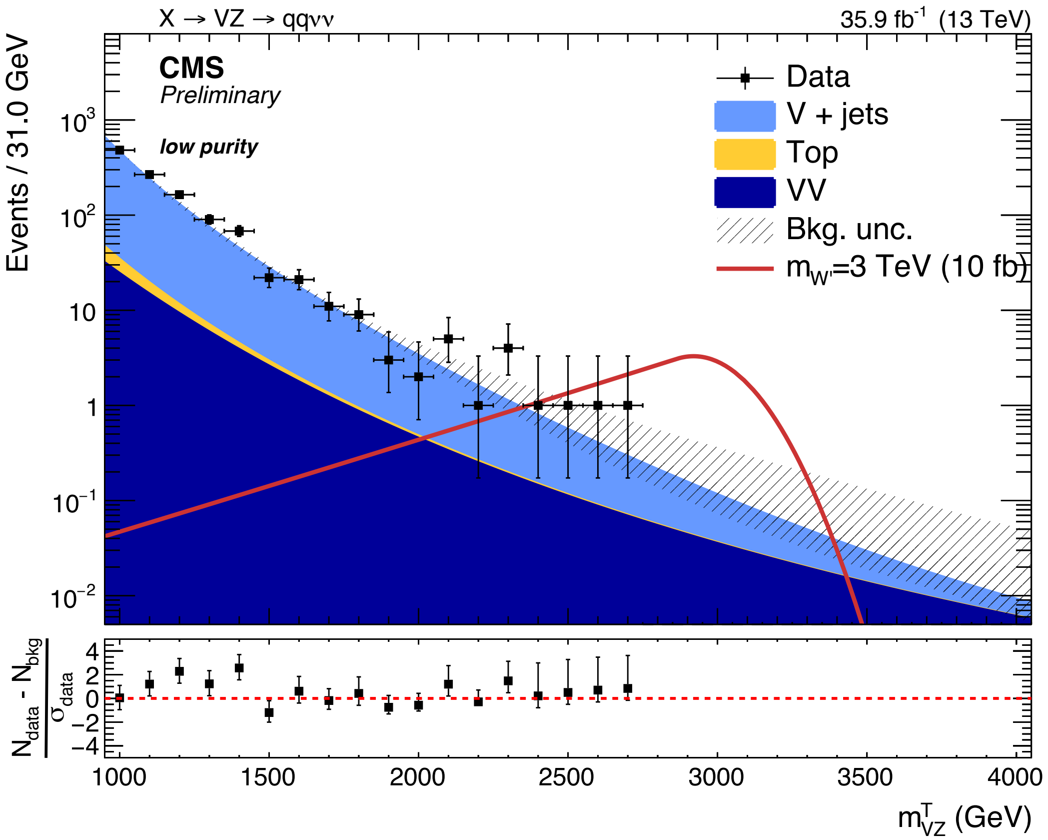


B2G-17-005

# alpha-ratio background estimation

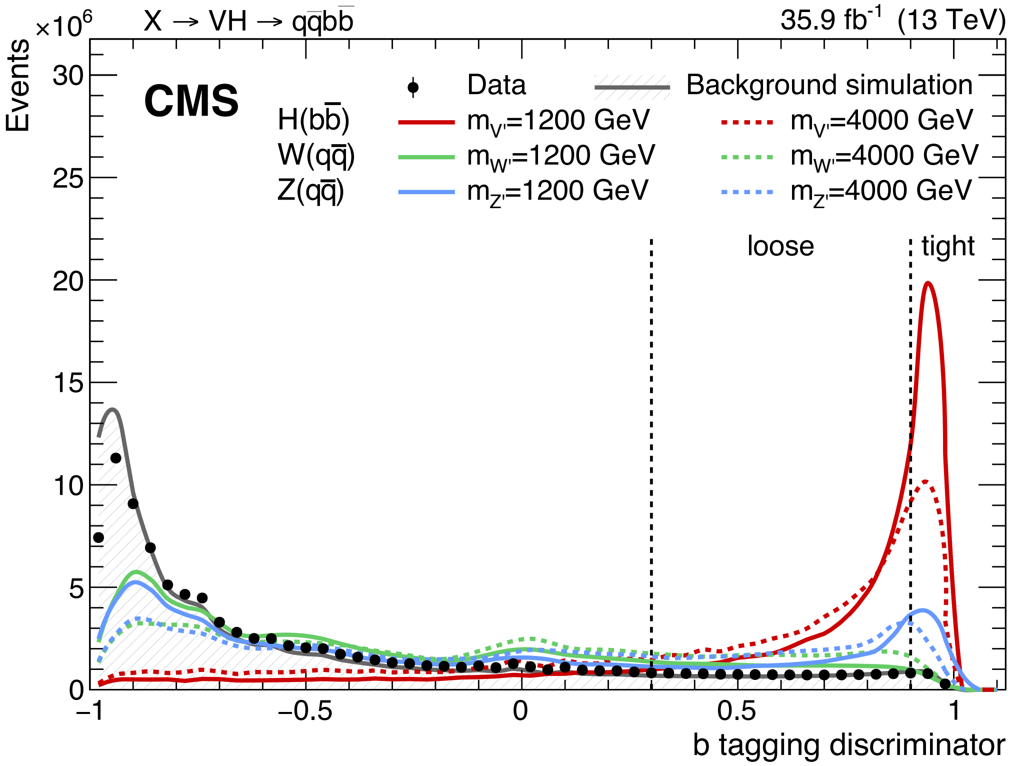
- Statistics in MC simulated samples still limited
- Furthermore, analysis performed in **extreme phase space**
- Use **jet mass sidebands** (40-65 GeV, 135-160 GeV) to exploit correlation between soft-drop jet mass and resonance mass
  - Higgs mass region kept blind
- Determine **ratio** of **simulated to data** distributions in sideband
- **Extrapolate** to signal region using **transfer function** (based on simulation)
- Method accounts for data-MC differences in shape and normalisation

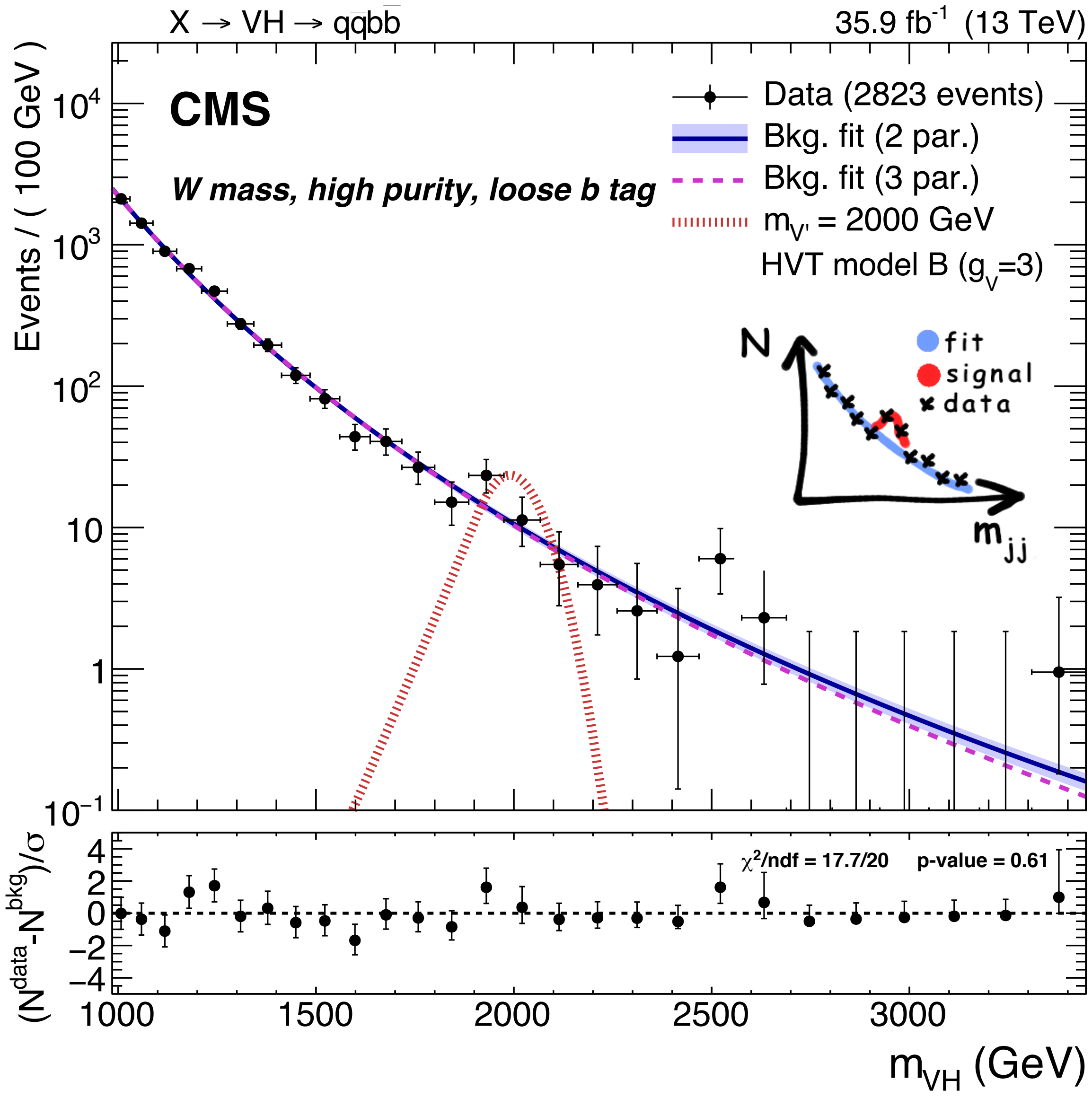


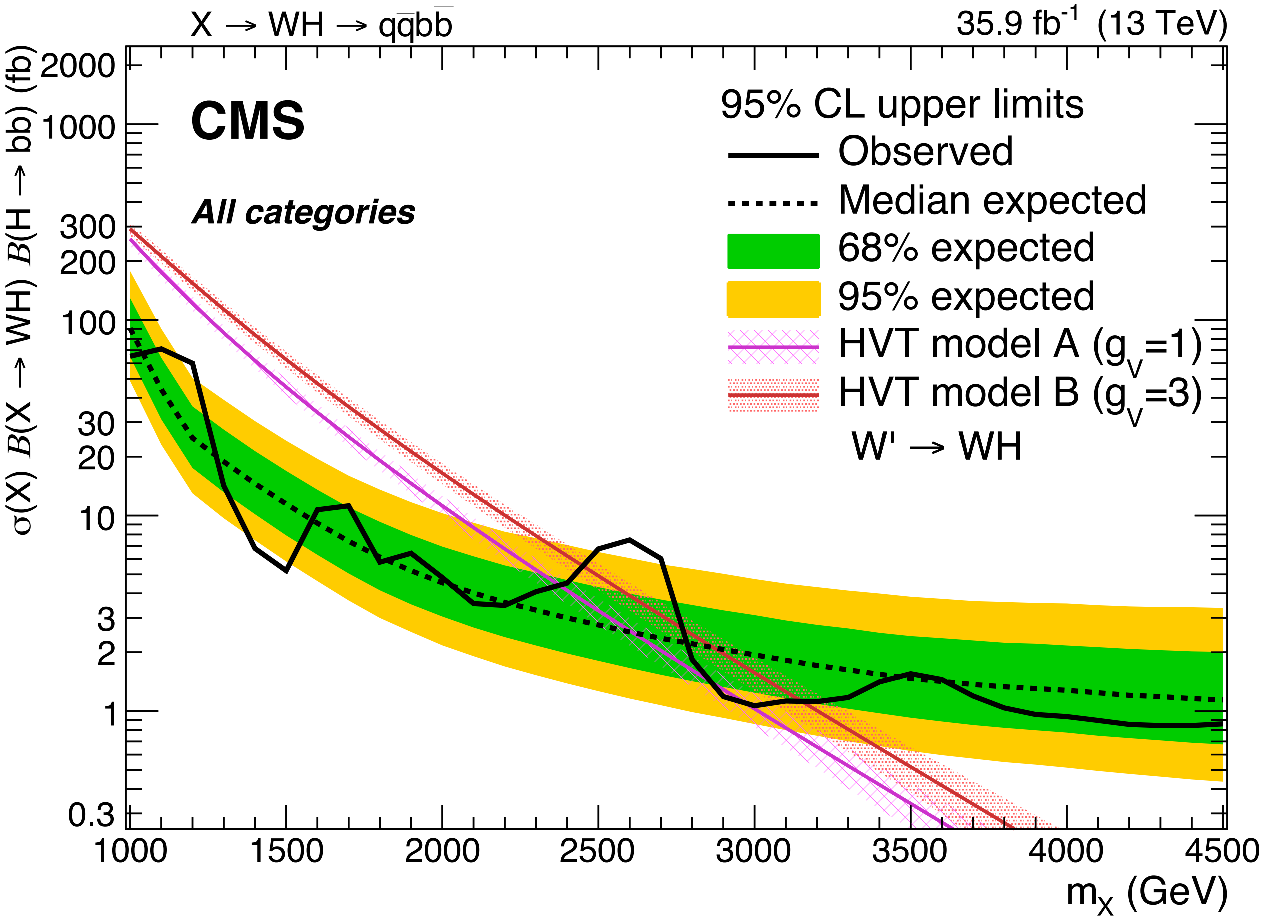


# Vh analysis

- Dijet final state: Wh, Zh, h➜bb
- Using dedicated Higgs tagger
  - Loose and tight regions
- Same background estimation strategy as for VV analysis



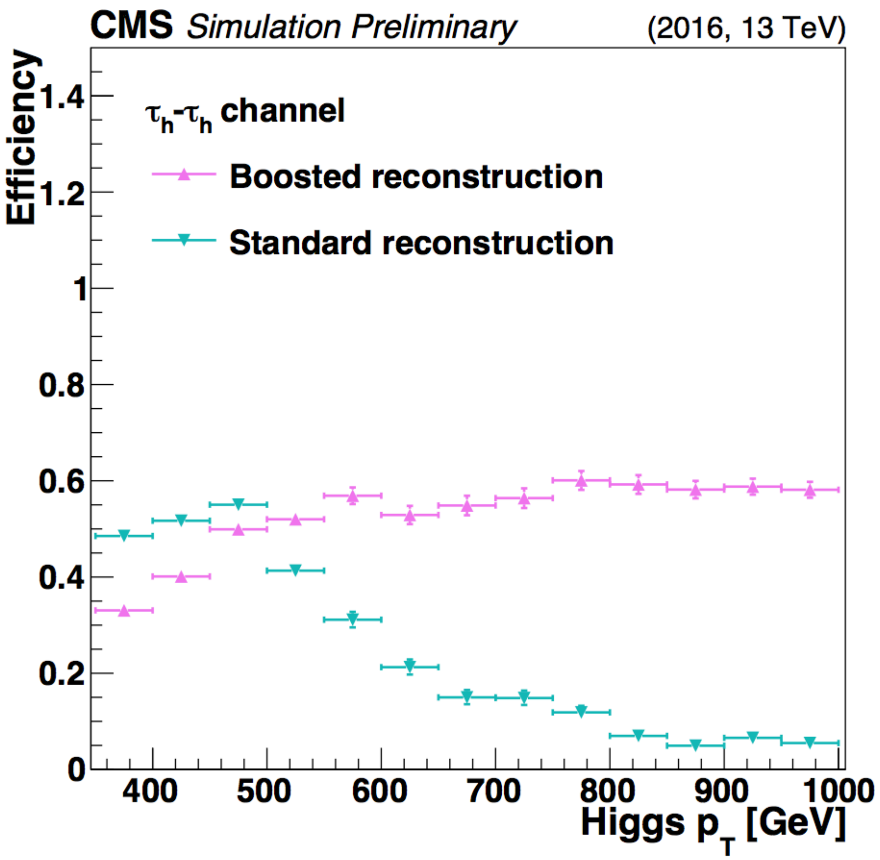


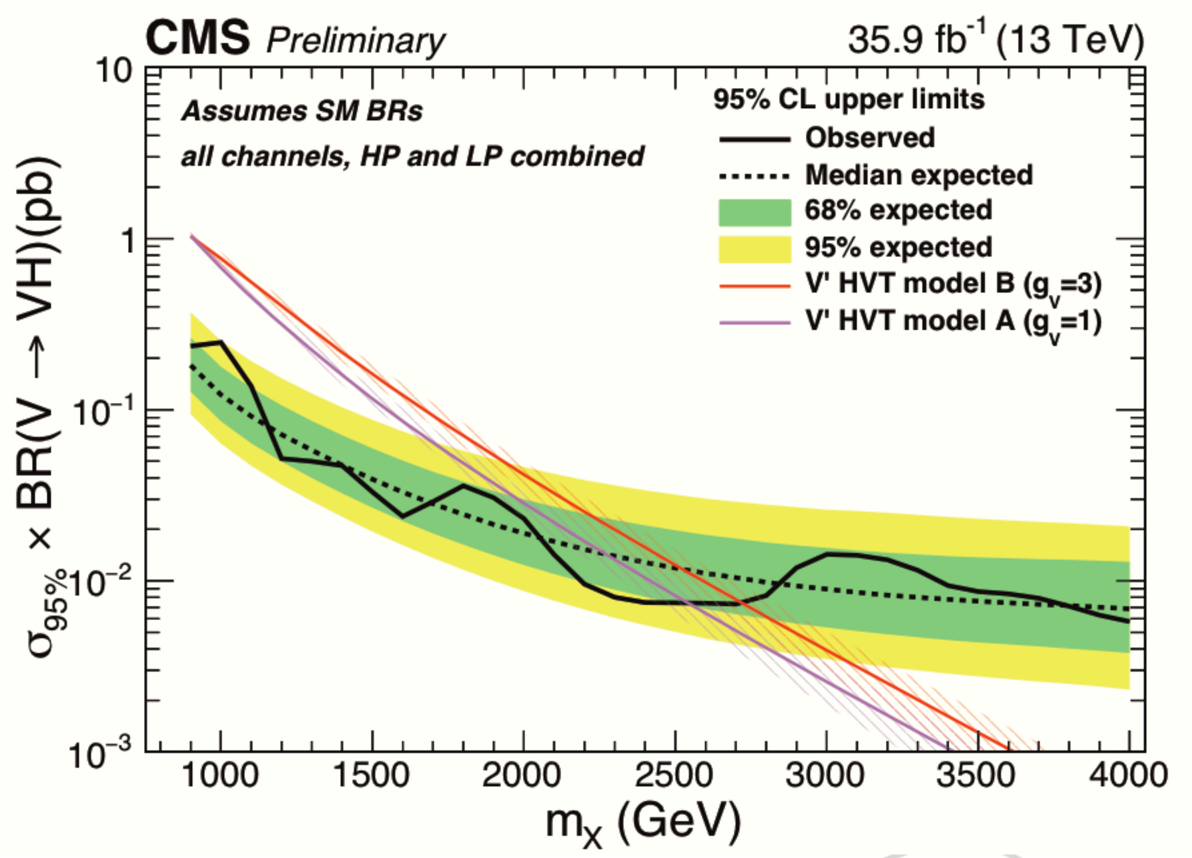


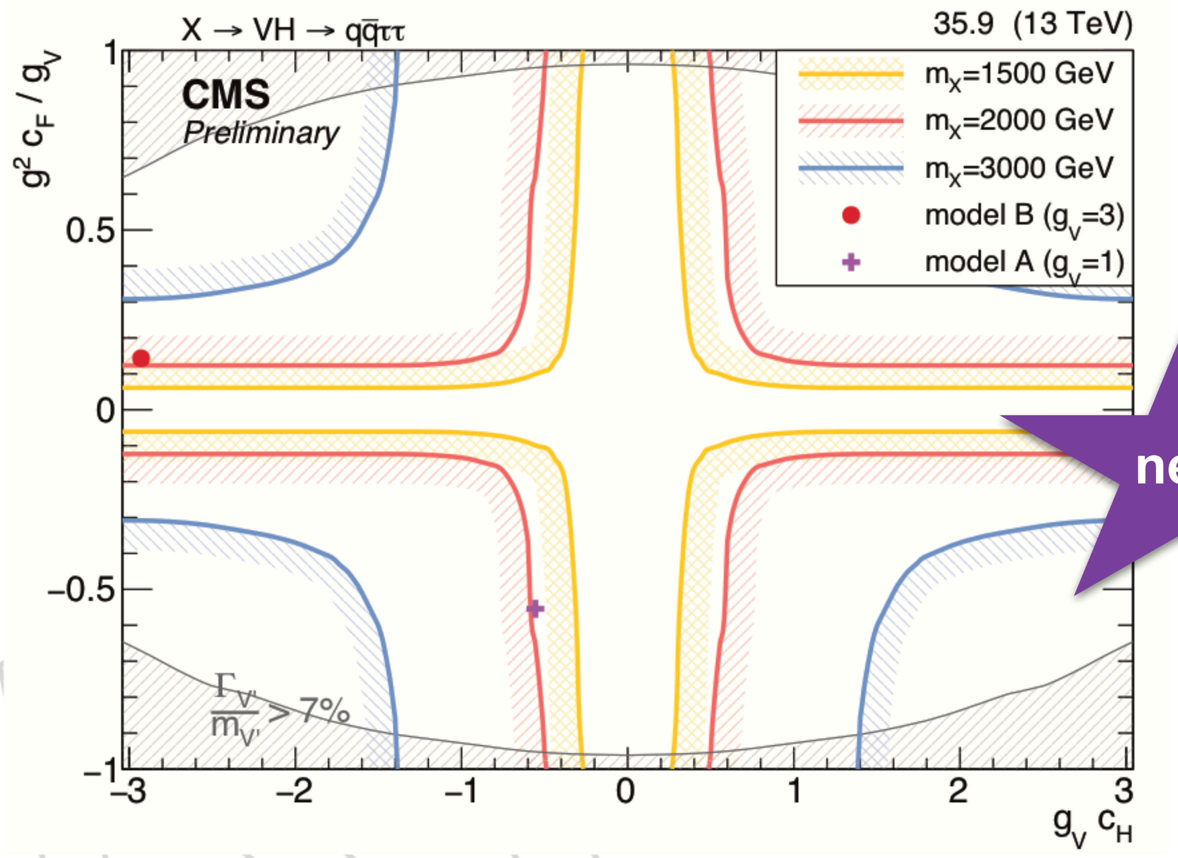
# VH/HH ➜ qq/bb ττ analysis

B2G-17-006
CMS-DP-2016-038

- Uses dedicated high-p$_T$ H➜ττ reconstruction
  - starting from fat jets, subjets reconstructed and then subject to τ reconstruction and identification algorithms
- Spin 0, 1, and 2 interpretation

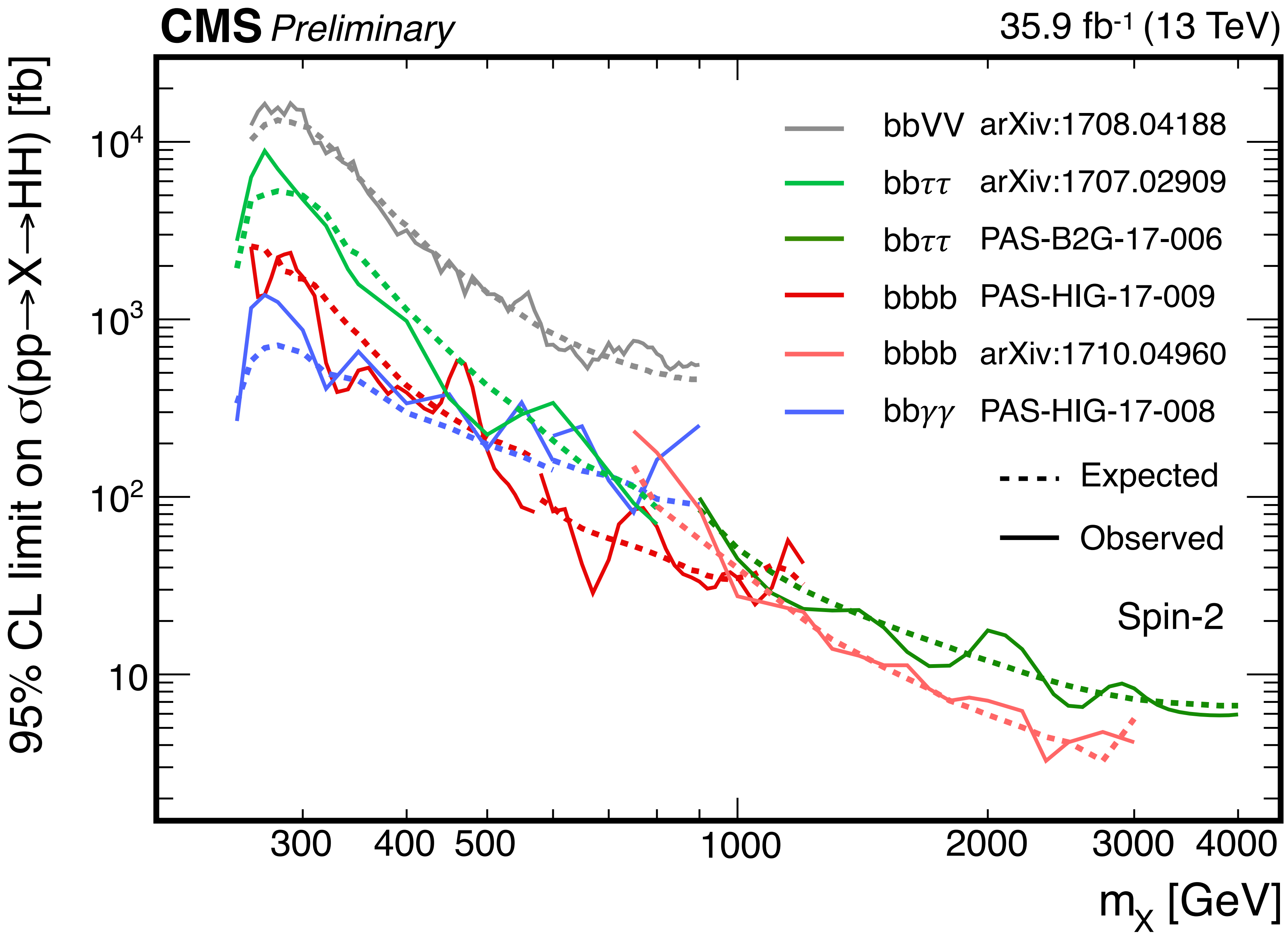






new!

# hh spin-2 interpretation



| | |
|---|---|
| bblνlν (HIG-17-006) | bbbb (HIG-17-009) |
| bbττ (HIG-17-002, submitted to Phys. Lett. B) | bbbb (B2G-16-026, submitted to Phys. Lett. B) |
| bbττ (B2G-17-006, submitted to J. High Energy Phys.) | bbγγ (HIG-17-008) |