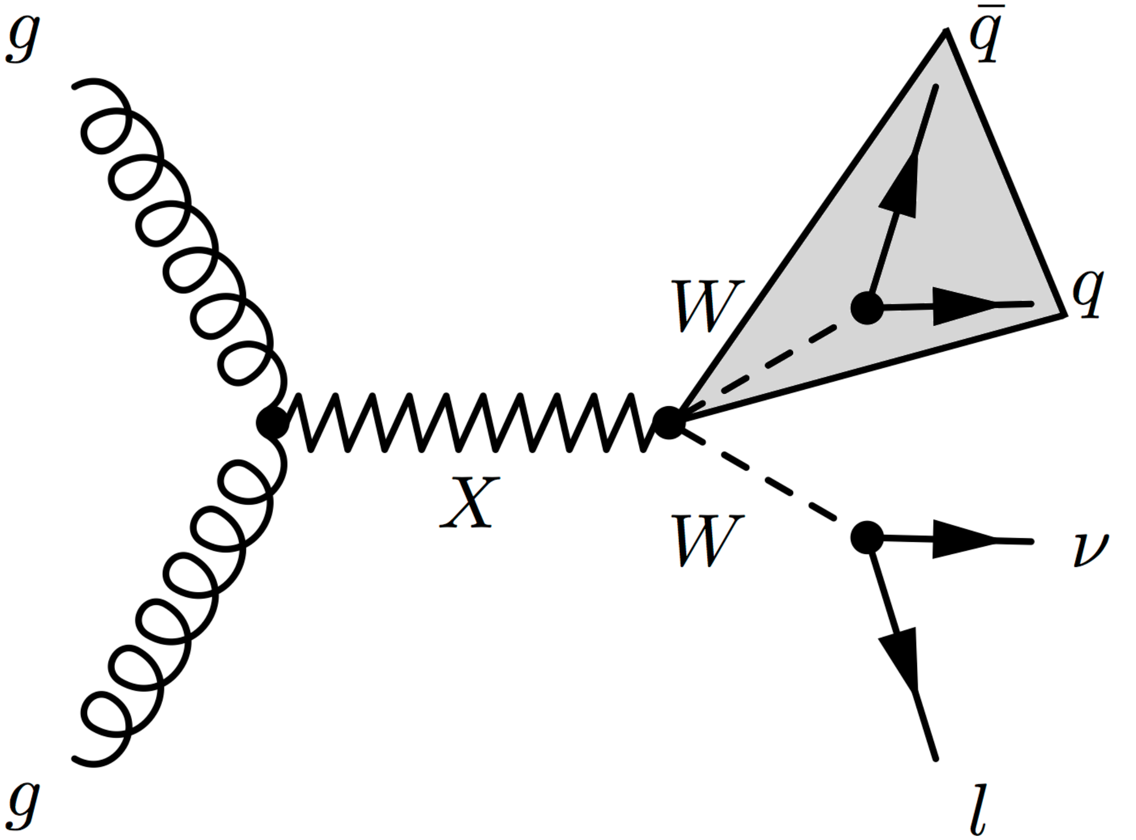
Search for diboson resonances at CMS

identifying highly energetic boson decays and discriminating new physics signals from the standard model background



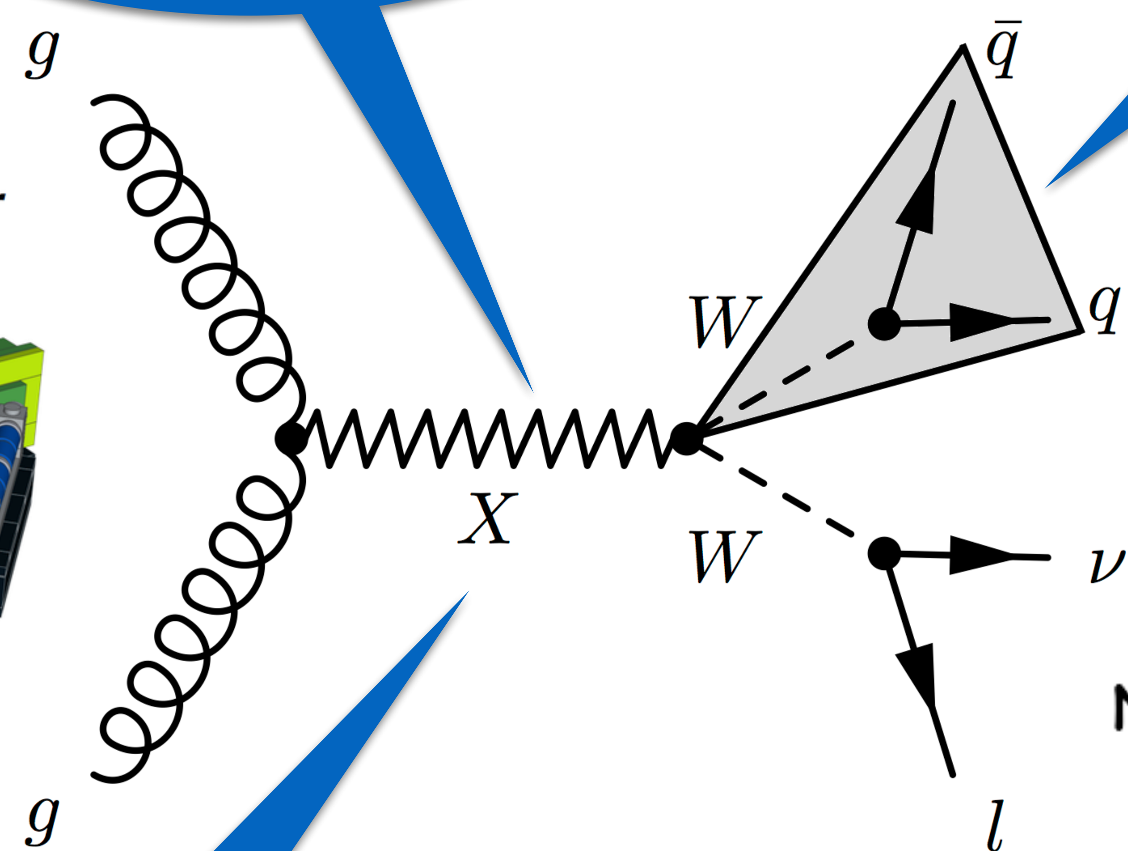
Clemens Lange (CERN)
DESY Physics Seminar
Hamburg/Zeuthen
31st May/1st June 2016

[https://ideas.lego.com/projects/94885]

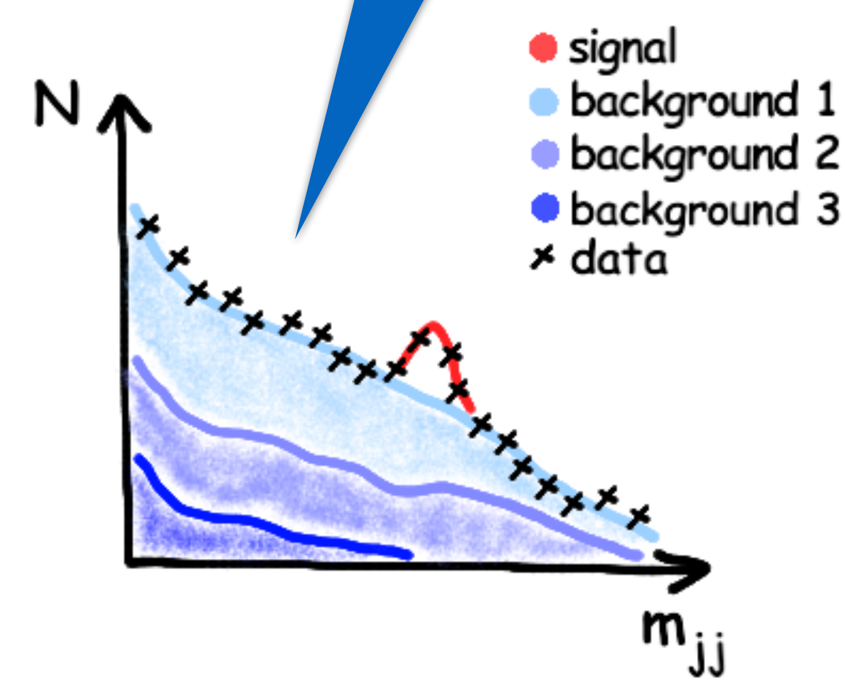
About diboson resonances

heavy resonances?

boson jets?



background estimation?



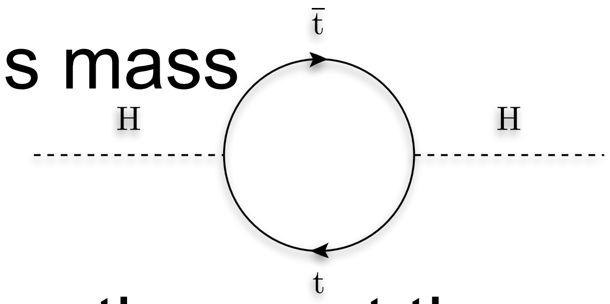
(wasn't there this diphoton resonance?)

We have a problem

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

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

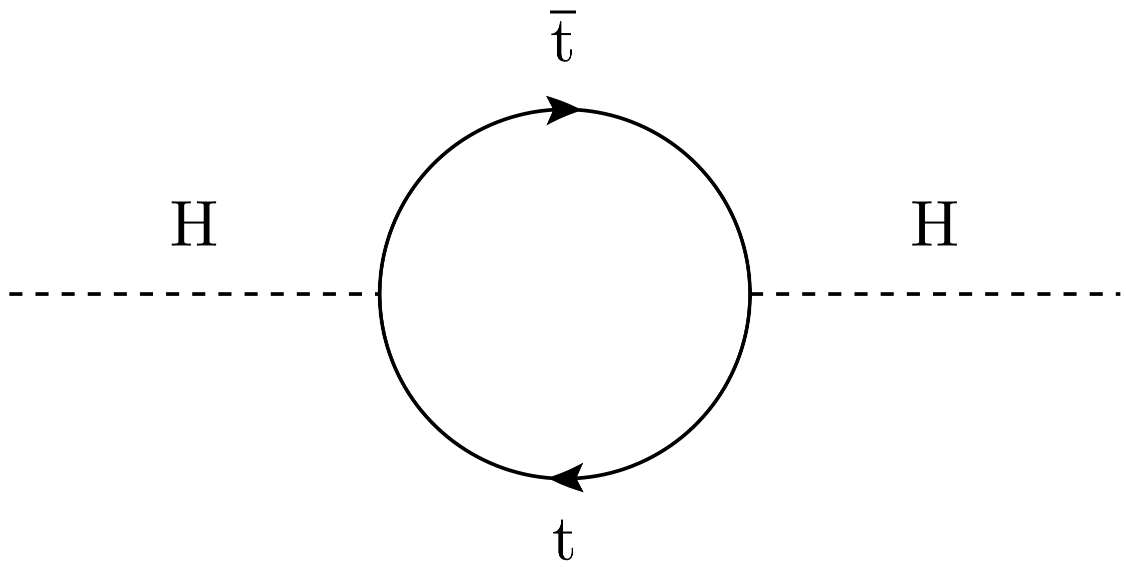
- no **symmetry** in the standard model (SM) protects the Higgs mass
- $\mu|H^2|$ always a singlet under phase transformations
- „natural“ explanation would be that SM is replaced by another theory at the TeV scale: $\mu^2 \sim (\text{heavier scale})^2$ ➜ **new particles**

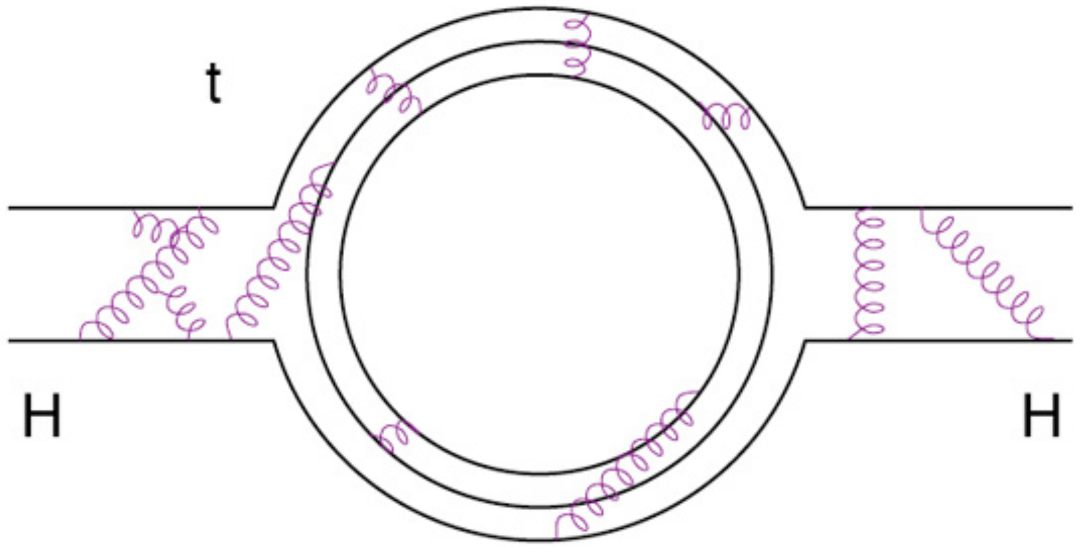


- these theories could be:
- **SUSY**: protecting the Higgs mass by a symmetry
- **Composite Higgs**: the Higgs is not elementary
- **Large/warped extra dimensions**: gravity is strong at electroweak scale

Composite Higgs?

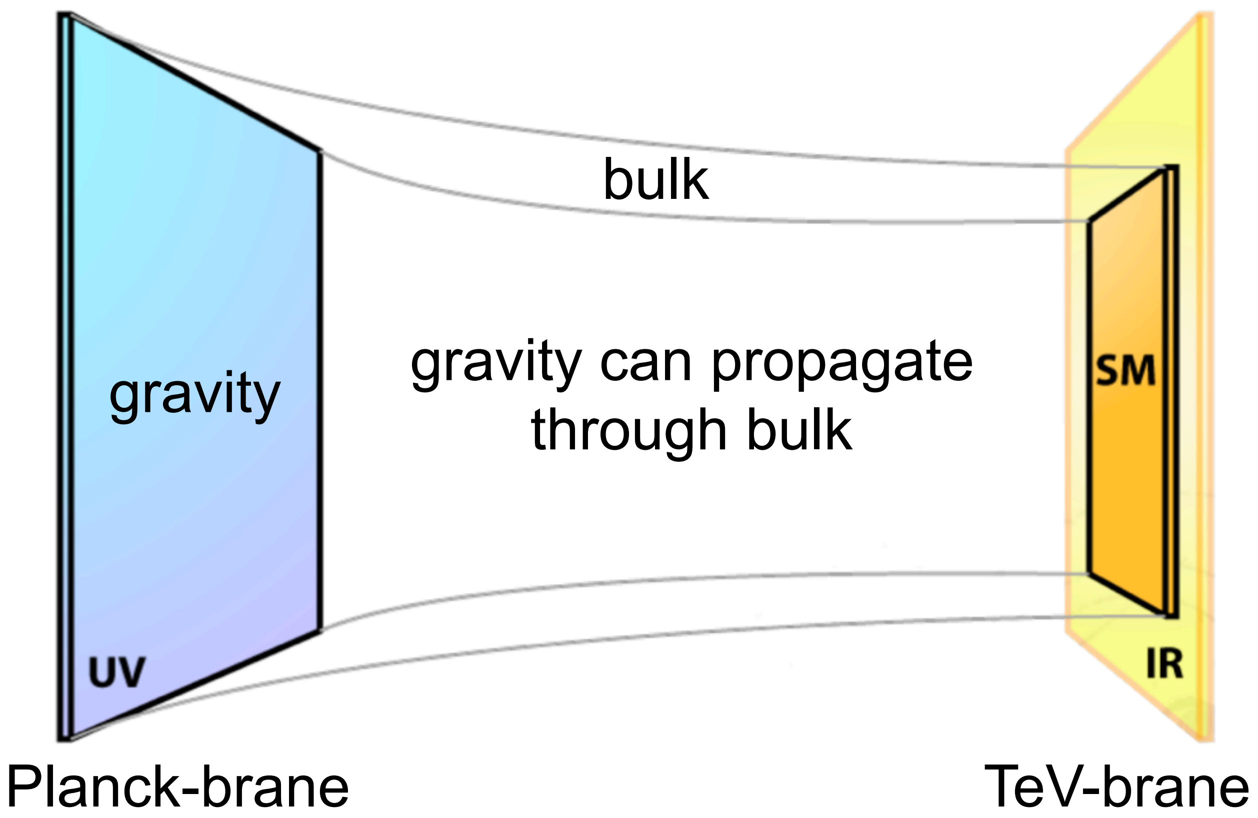
- the **Higgs** could be **non-fundamental**
- instead: **bound state** of a new strong interaction
- e.g. size of 10^{-18} m ~ **Fermi scale** (100 GeV)
  - light Higgs like a pion from a new sector
- solves hierarchy problem, and brings along **new heavy particles**/states
- heavy partners of SM particles decay to lighter ones (W, Z, H, top, …)





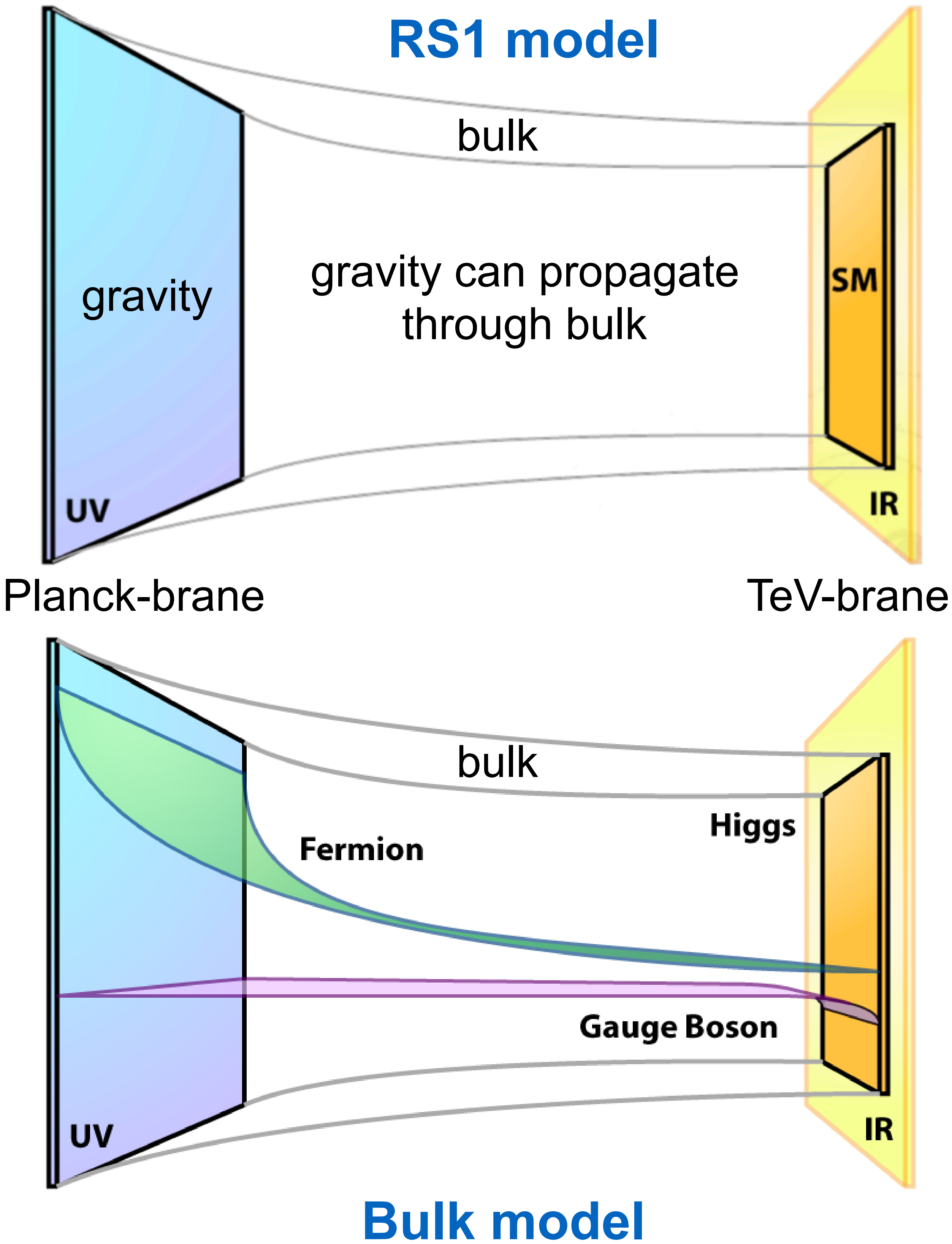
Large extra dimensions?

- another attempt to solve the hierarchy problem
- SM fields are confined to **four-dimensional „membrane“**, gravity propagates in additional dimensions
- effectively, **change power law of gravity** from $1/r^2$ to $1/r^{2+N}$, where N = number of extra dimensions
- this only applies to particles with $r \ll N$ - **smaller things have more possibilities to move**
- „large“, because of size 1 mm to ~1/TeV
- proposed by Arkani-Hamed, Dimopoulos, and Dvali (ADD)



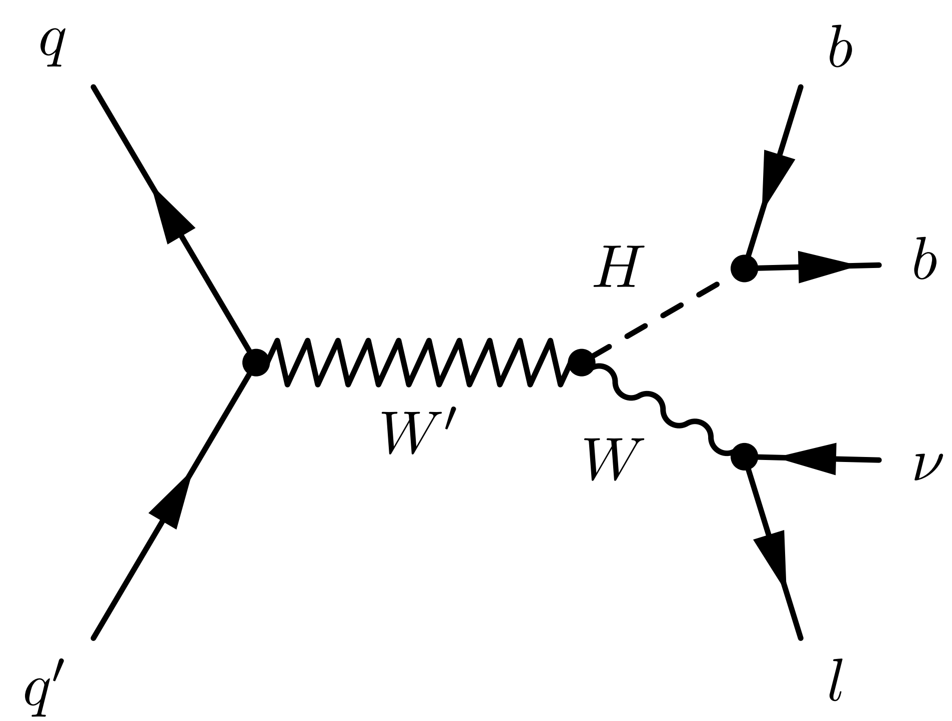
Warped extra dimensions?

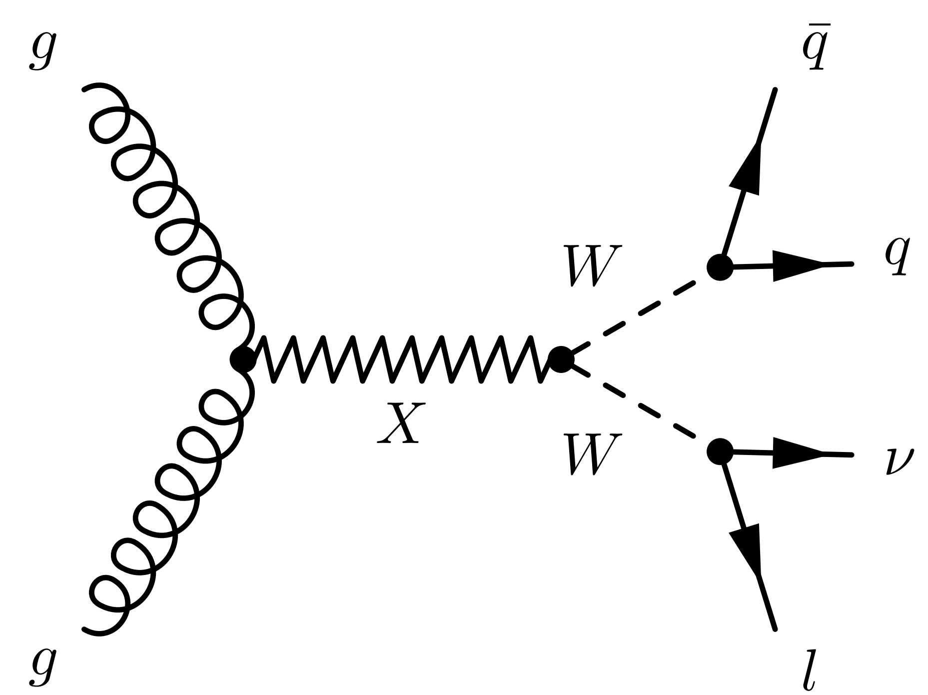
- often referred to as **Randall-Sundrum** (RS) models
- warping causes energy scale at one end of the extra dimension to be much larger than at the other end
- SM models reside on **TeV-brane** (in RS1 models)
- **bulk graviton** models allow SM particles into 5D-bulk
- overlap of 5-D profiles at TeV-brane (and Higgs) determine particle masses
- additionally, if distance between two branes is not fixed, additional fluctuations can occur



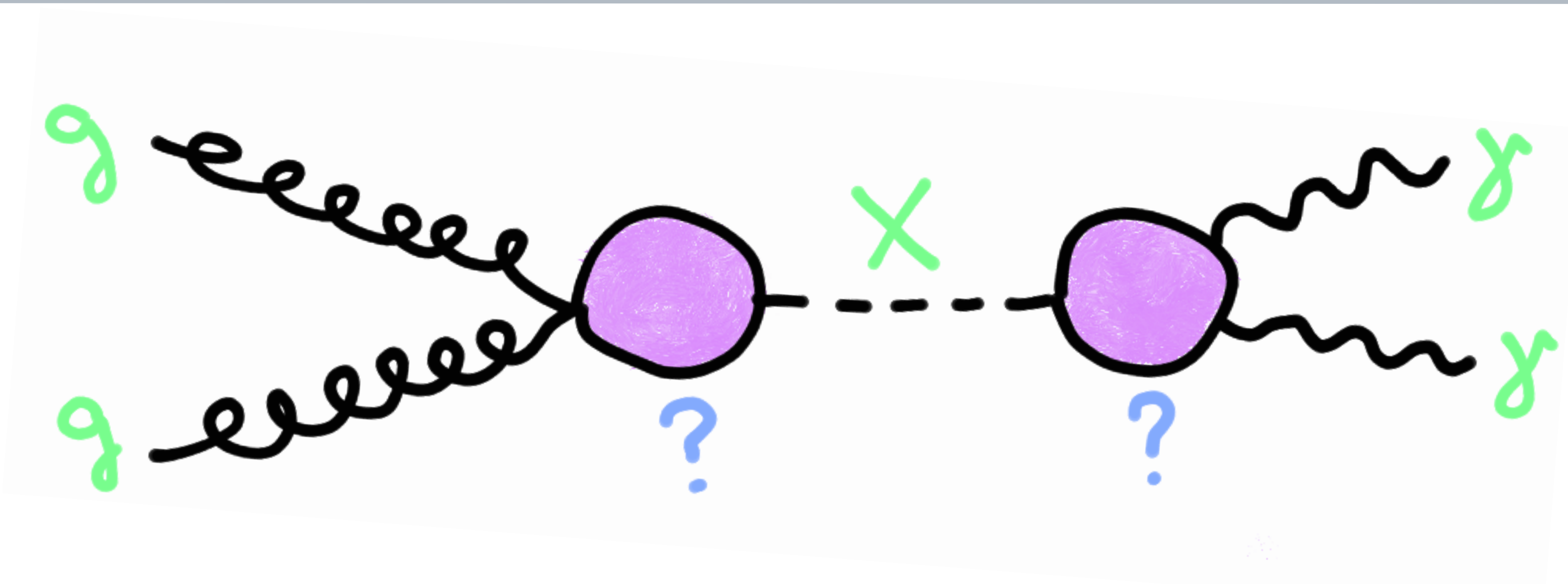
How do we observe these models at the LHC?

- should be able to observe excitations/ resonances/fluctuations
- **composite Higgs**: electroweak composite vector resonances
  - mostly spin-1 (W', Z')
  - decay to pairs of W, Z, H
- **Randall-Sundrum**: Kaluza-Klein excitations of gravitons + radion fluctuations
- gravitons (spin-2):
  - RS1: decay predominantly to leptons
  - bulk: decay to pairs of W, Z
  - ADD: broader excess from many narrow-spaced resonances
- radions (spin-0):
  - only used for signal modelling here
- focus here on **narrow resonances** (width < detector resolution), **mass ≥ 600 GeV**





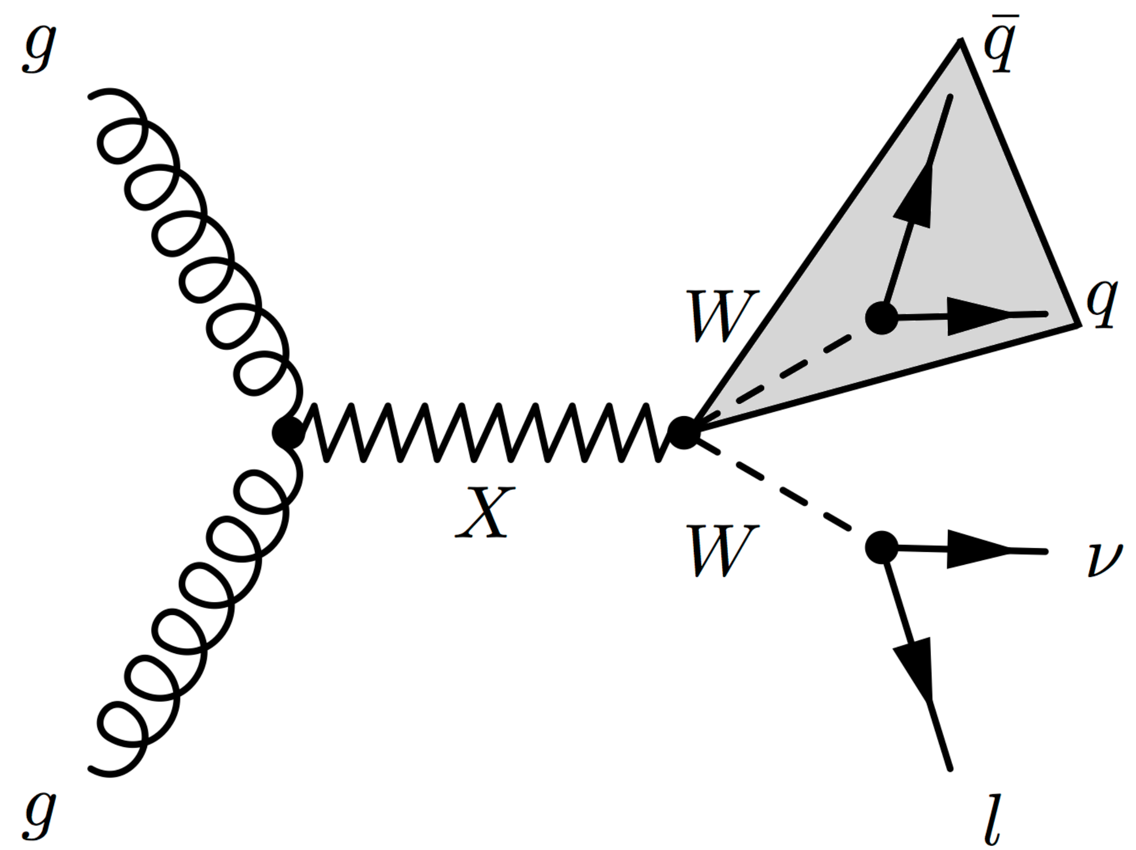
What about the diphoton resonance?



- neutral resonance could be **graviton** or **radion** as in diboson searches
- resonance cannot directly couple to photons ➜ loop of **charged particles** (e.g. W, top, ?) **in decay** (and production?)
- there must be **more than just a di-photon resonance**
- searches presented in this talk **constrain** what **physics models** this potential resonance could be

Reconstructing heavy resonances

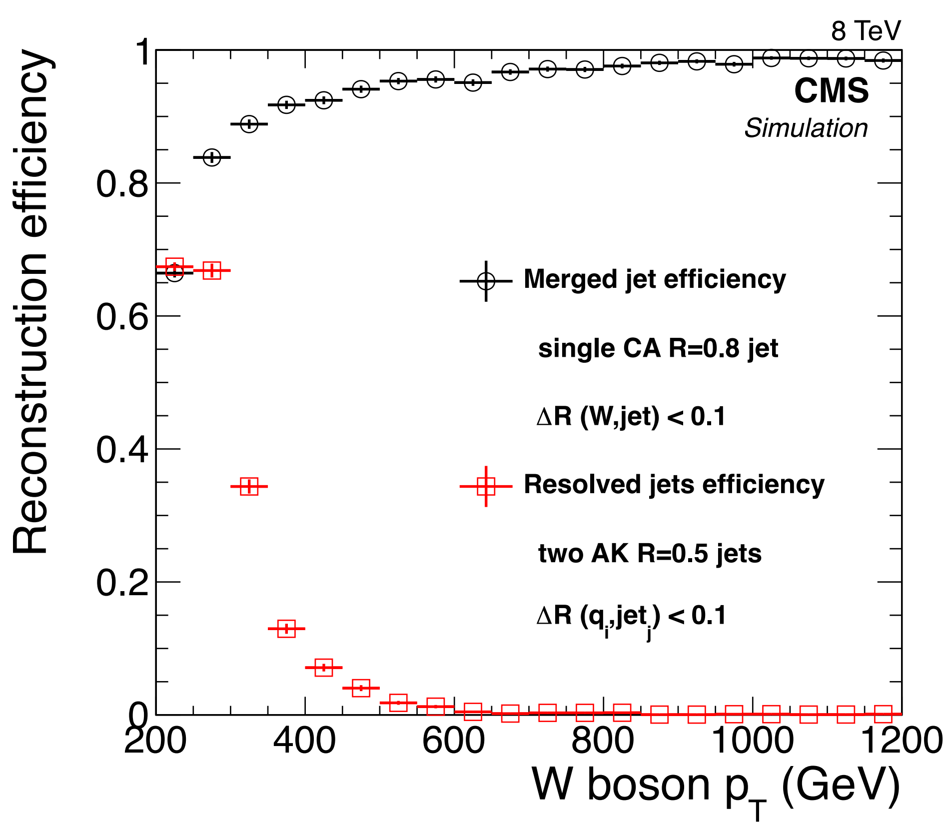
- bosons will be very energetic ➜ **collimated decay products**
- need to develop dedicated reconstruction methods
- hadronic decays of bosons:
  - „**boson-tagging**“
  - exploiting substructure of jets
- leptonic decays:
  - special **isolation** for dileptonic decays
  - dedicated reconstruction algorithms for high-p_T leptons
  - new tau-identification algorithms



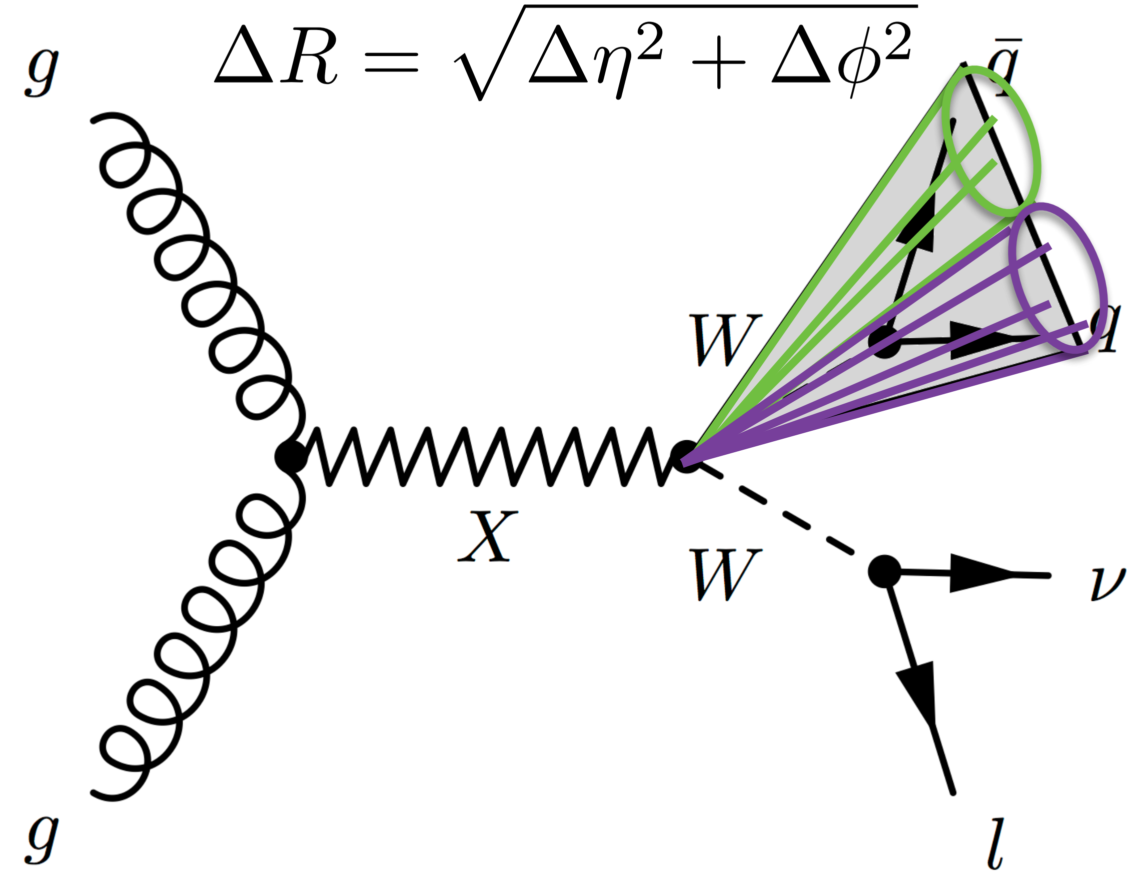
focus here on Run-2 developments and analyses

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**



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

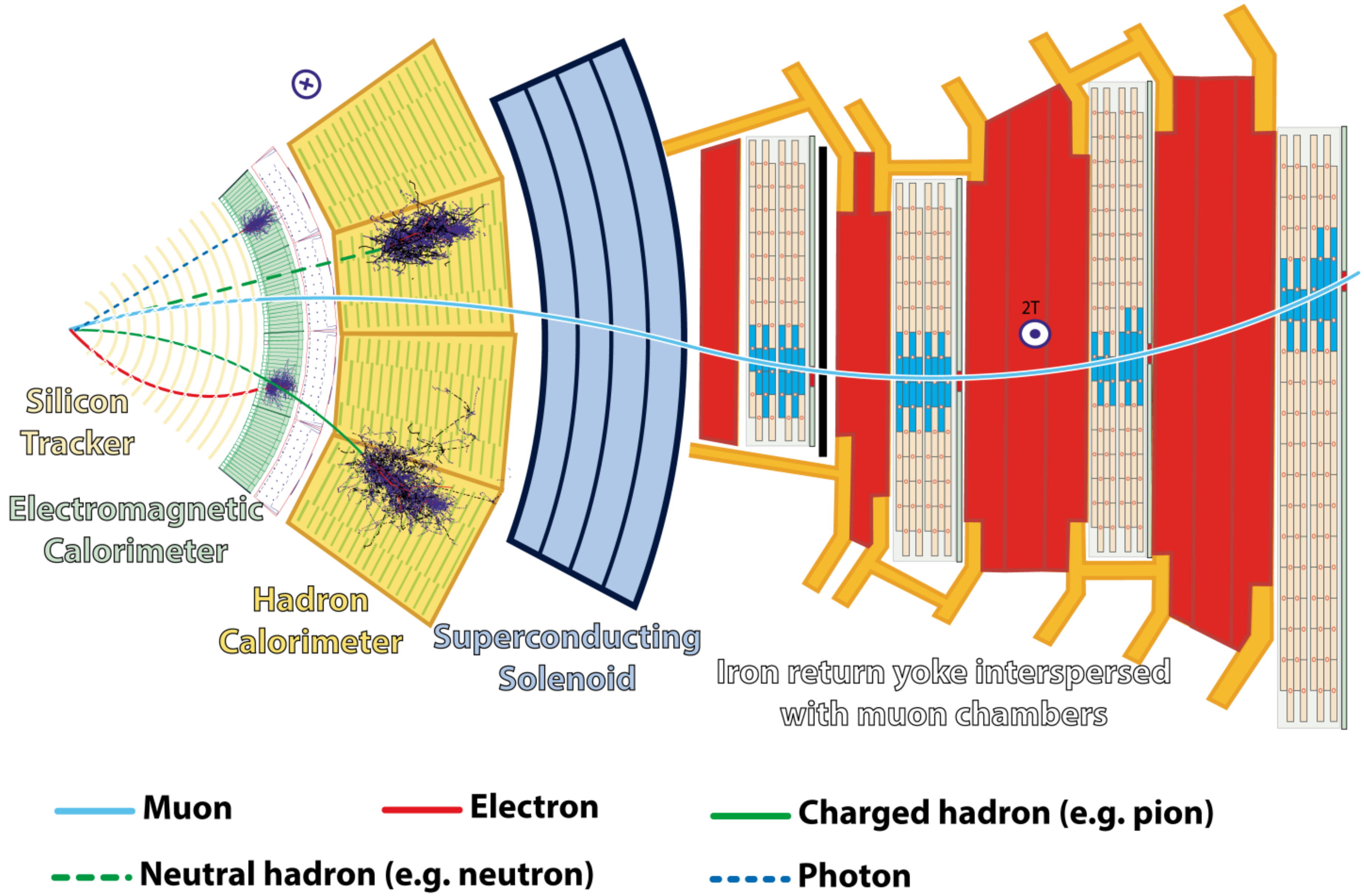


$$\Delta R_{qq} \approx 2\frac{M_V}{p_T^V}$$

back of the envelope calculation:

for a resonance of mass 1 TeV the bosons from the decay will have $p_T \sim 0.4$ TeV ➜ $\Delta R \approx 0.4$

CMS particle flow reconstruction

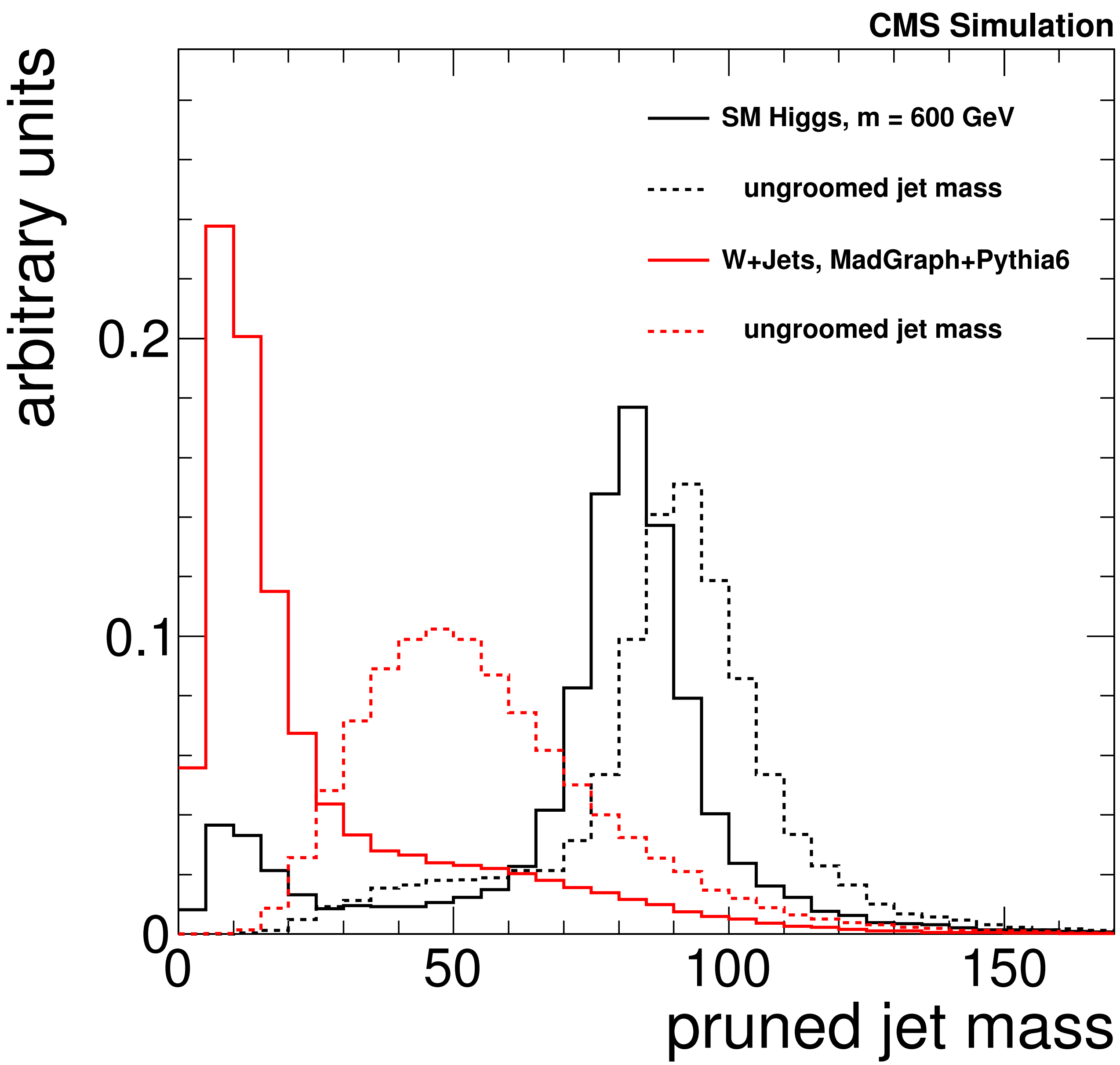


- tracking detectors and calorimeters contained in **magnetic field**
- particle flow algorithm makes use of sub-detectors with **best resolution** (both spatial and energy)
- actual „**particles**“ enter jet clustering

Jet pruning

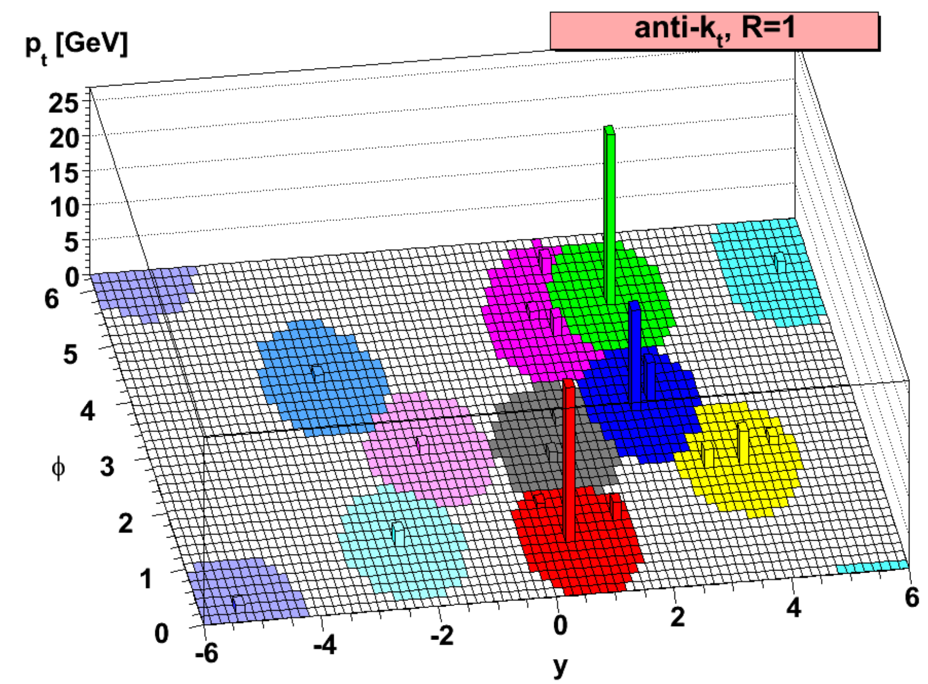
- we know the **masses** of W, Z and Higgs very well ➜ can use them as **constraints**
- however, large number of particles in jet ➜ rather bad resolution
- jet pruning (generally grooming) removes **soft and large angle radiation**
- strategy:
  - recluster jet using Cambridge-Aachen (CA) jet algorithm

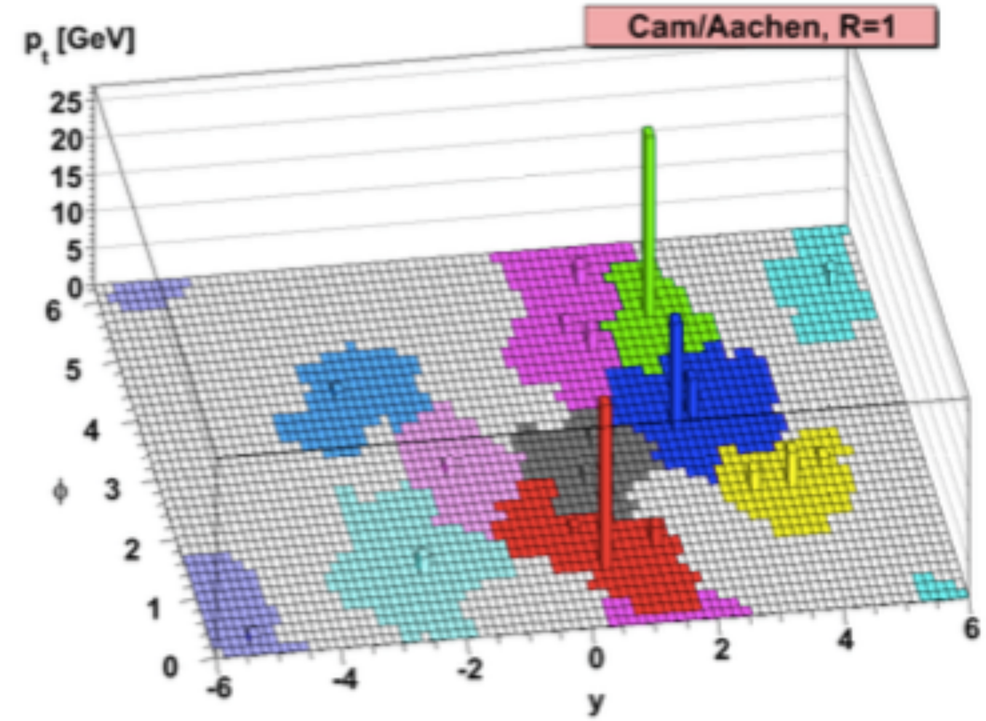
?



Reminder: jet clustering algorithms

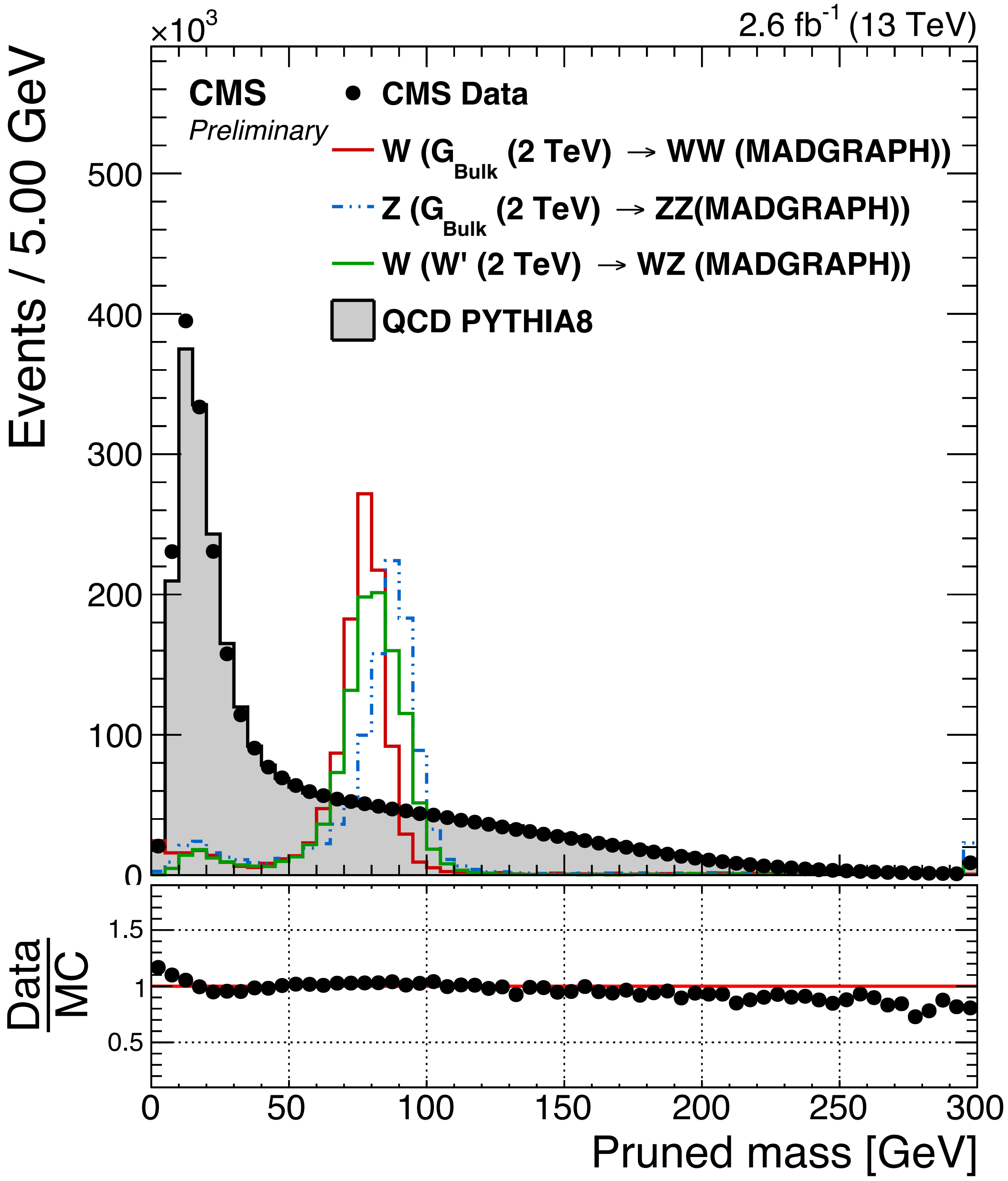
- k_T-algorithms: **sequential clustering**
- examine four-vector inputs pairwise and construct jets hierarchically
- **anti-k_T**: preferentially merge constituents with **high p_T** with respect to their nearest neighbours first
- **Cambridge-Aachen**: no pT-weighting, merge based on **spatial separation**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





Jet pruning

- 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 pruning (generally grooming) removes **soft and large angle radiation**
- strategy:
  - recluster jet using Cambridge-Aachen (CA) jet algorithm
  - „soft“: $\min(p_T^i, p_T^j)/\tilde{p}_T < 1$
  - „large angle“: $\Delta R_{ij} > m^{\text{orig}}/p_T^{\text{orig}}$, orig = unpruned CA jet
- cut on mass window (~±10 GeV)



N-subjettiness

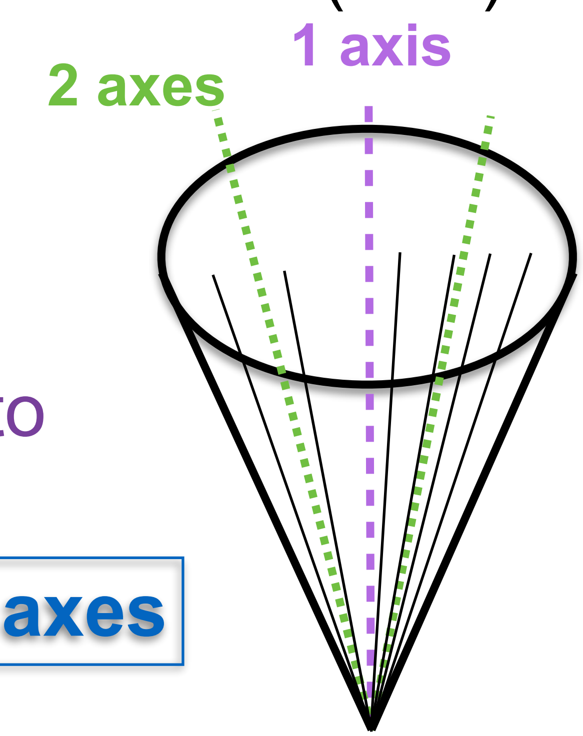
- for **boson-tagging**: want to quantify how **2-subjetty** a jet is
- ➜ to what extent is energy flow aligned along 2 momentum directions (N=2)?

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

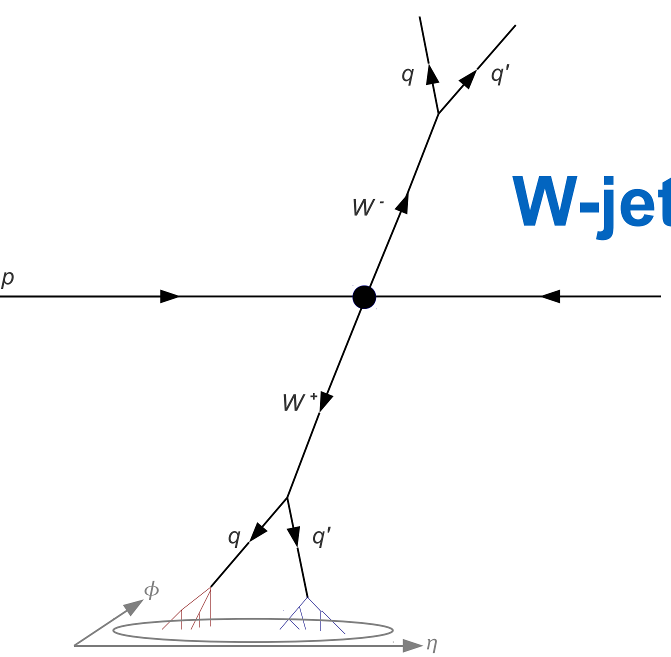
normalisation

sum over particles

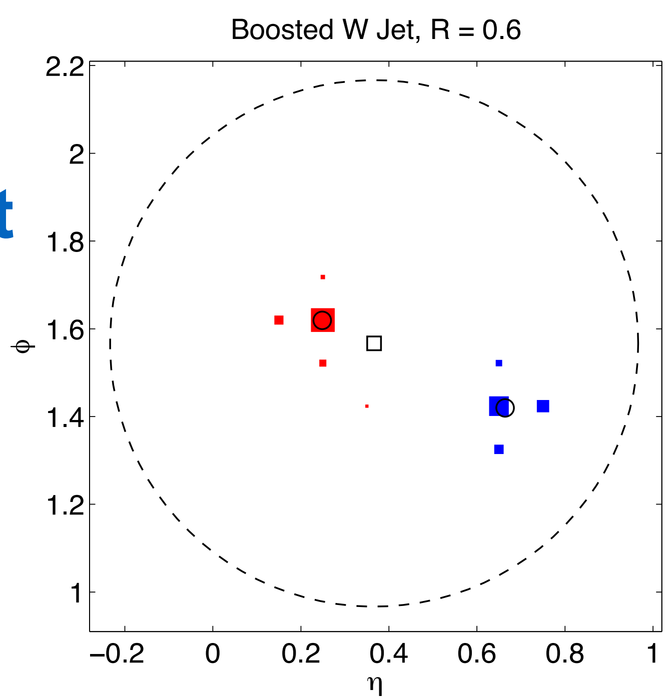
minimise distance to candidate subjets

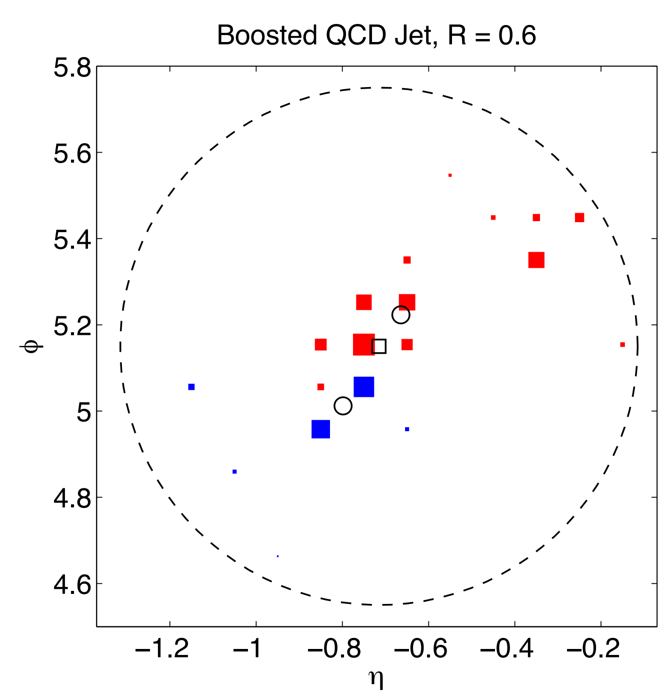


low values of τ_N ➜ compatibility with the hypothesis of N axes

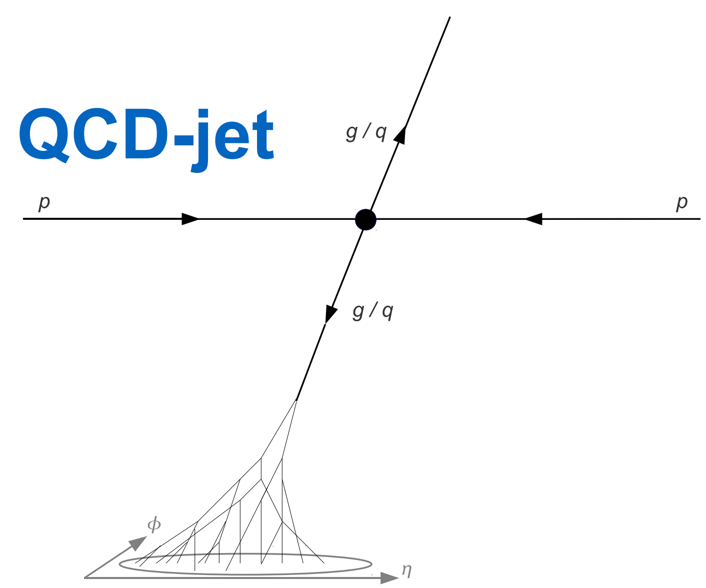


W-jet



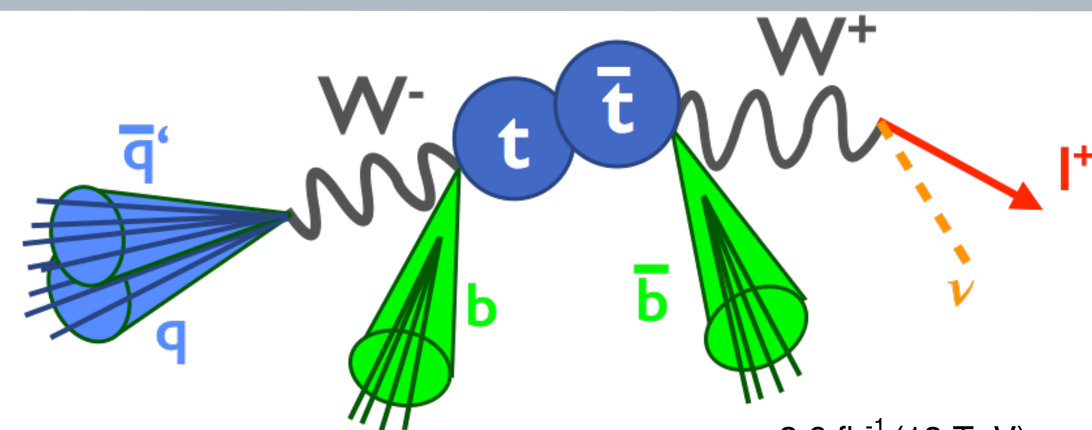


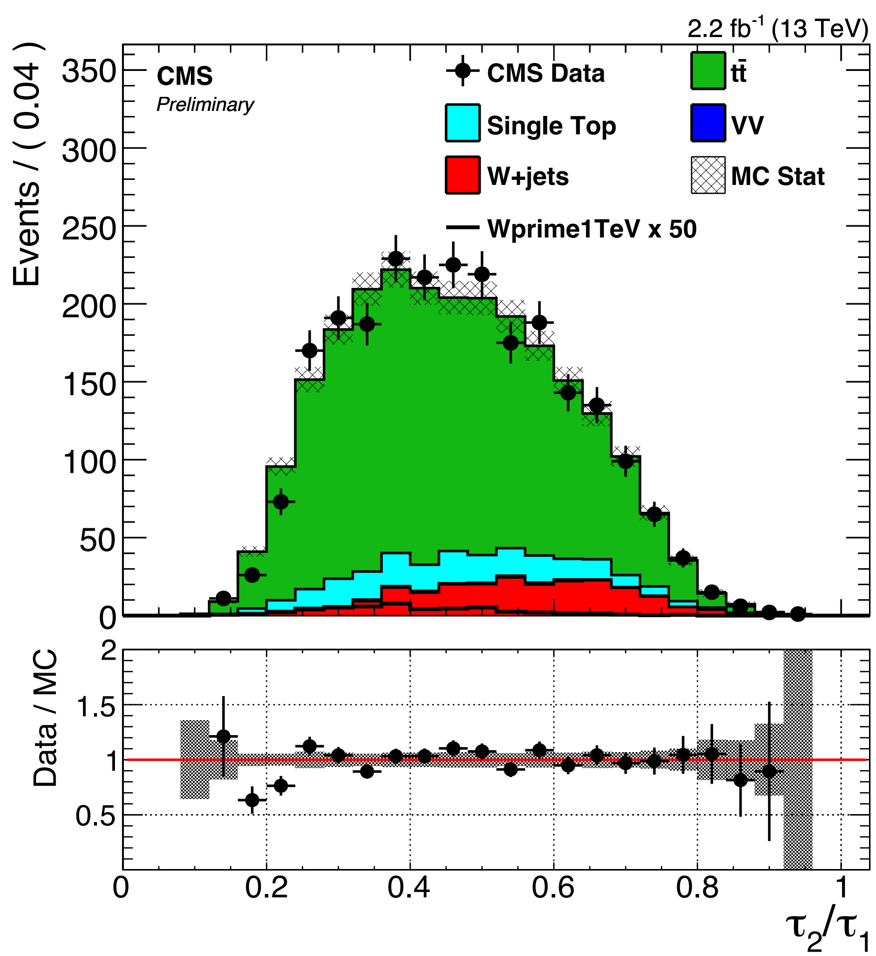
QCD-jet

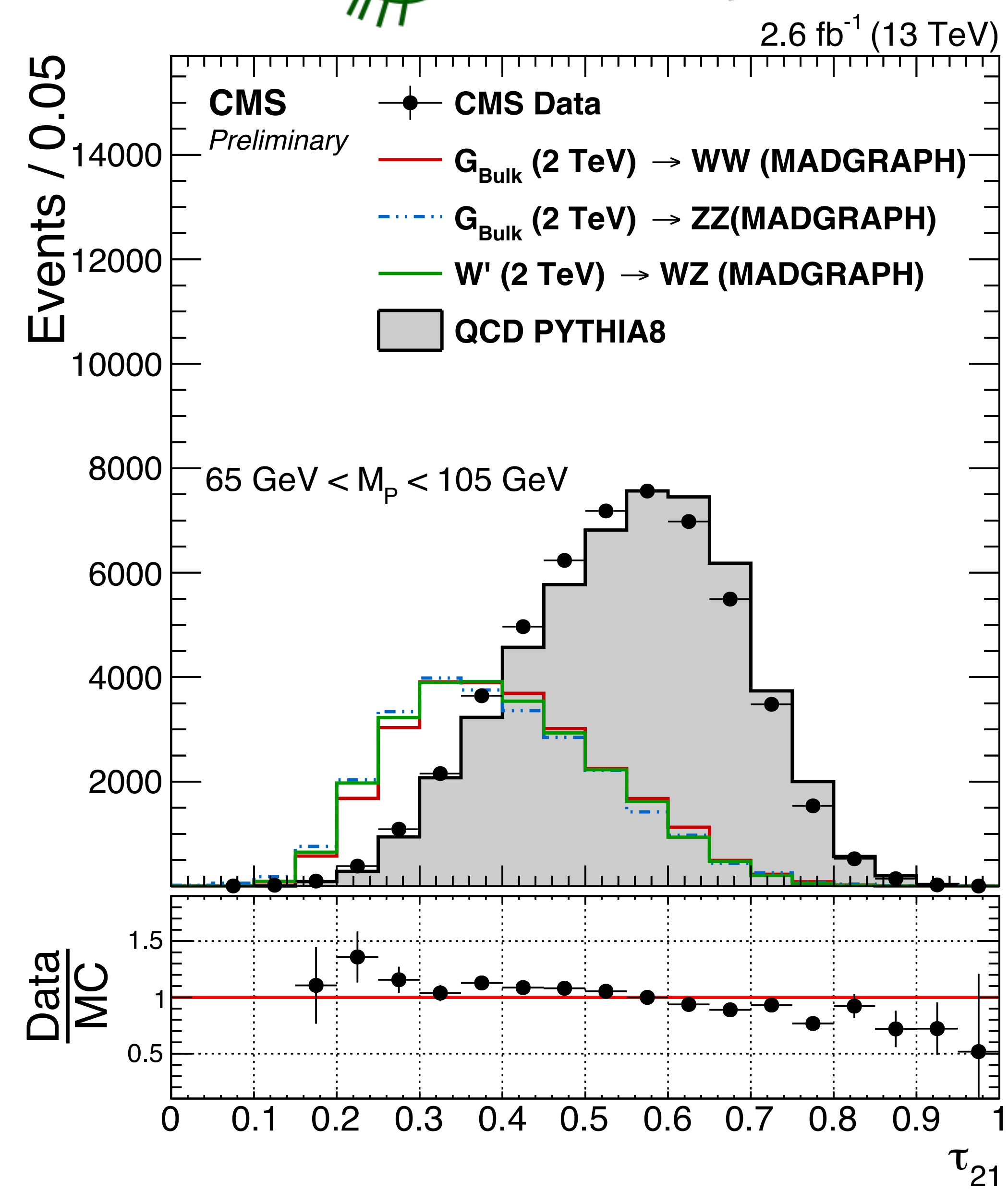


N-subjettiness ratio

- bare τ_N has very little discrimination power
- take **ratio τ_2/τ_1** instead
- mind: rather **complicated variable**, difficult to model ➜ need to validate in data
- clean sample of W-jets: **top-antitop quark pairs** used for calibration



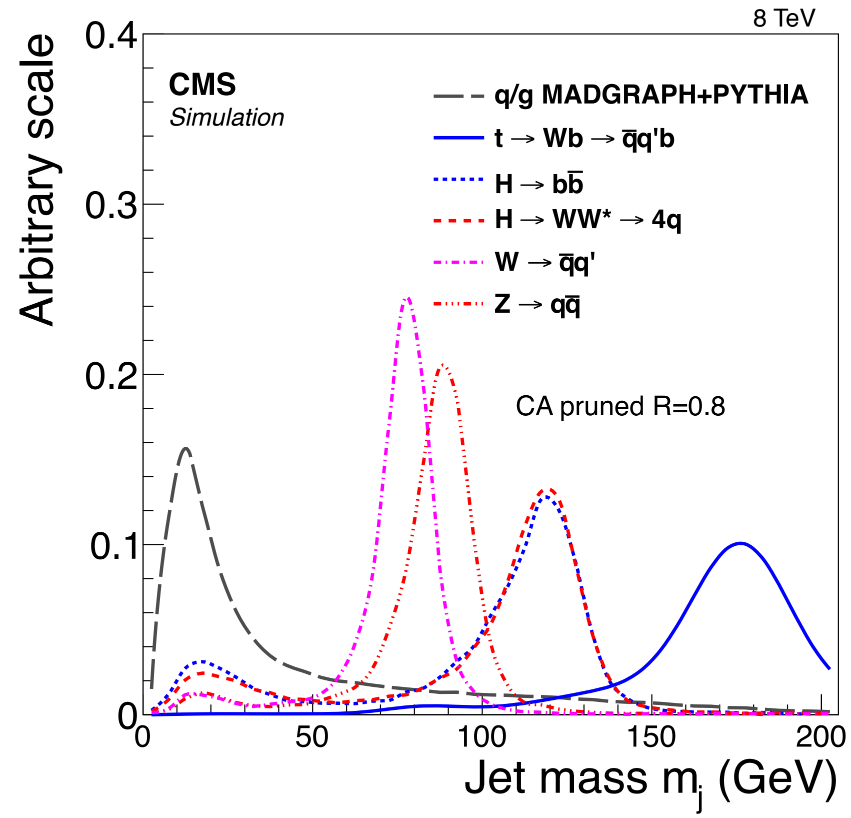


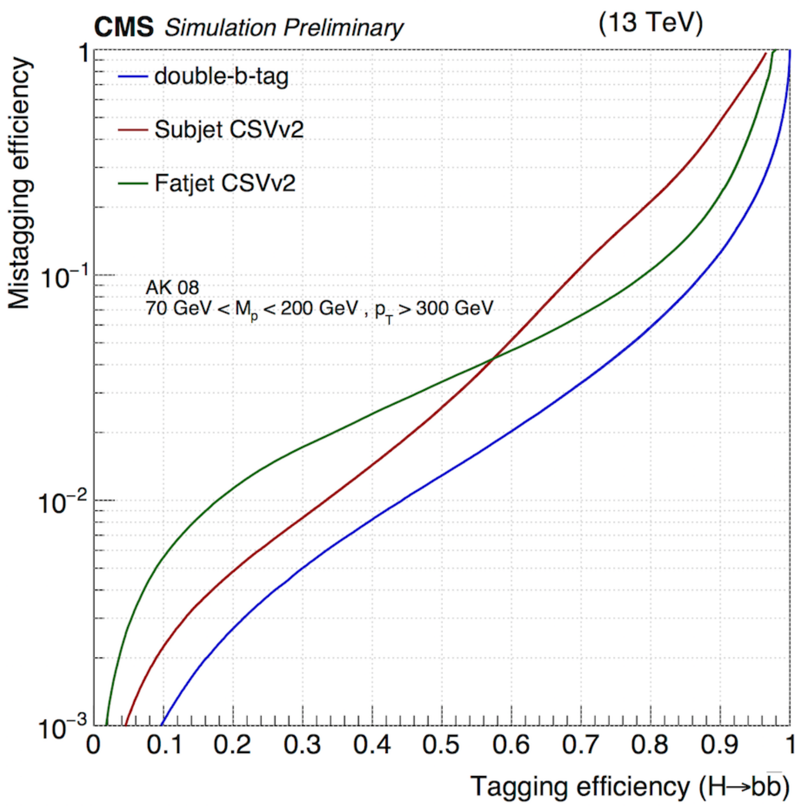


Higgs➜bb tagging

- Higgs has higher mass than W/Z bosons ➜ τ_2/τ_1 less important, **exploit b-jet content** instead
- two different strategies:
  - identify b-subjets
  - tag fat jet
- currently, both show comparable performance
- **50% lower mis-tagging rate** than W-/Z-tagging
- dedicated Higgs-tagger available soon

H➜WW➜qqqq tagging is done using τ_4/τ_2 ratio (cf. τ_2/τ_1 for W/Z)

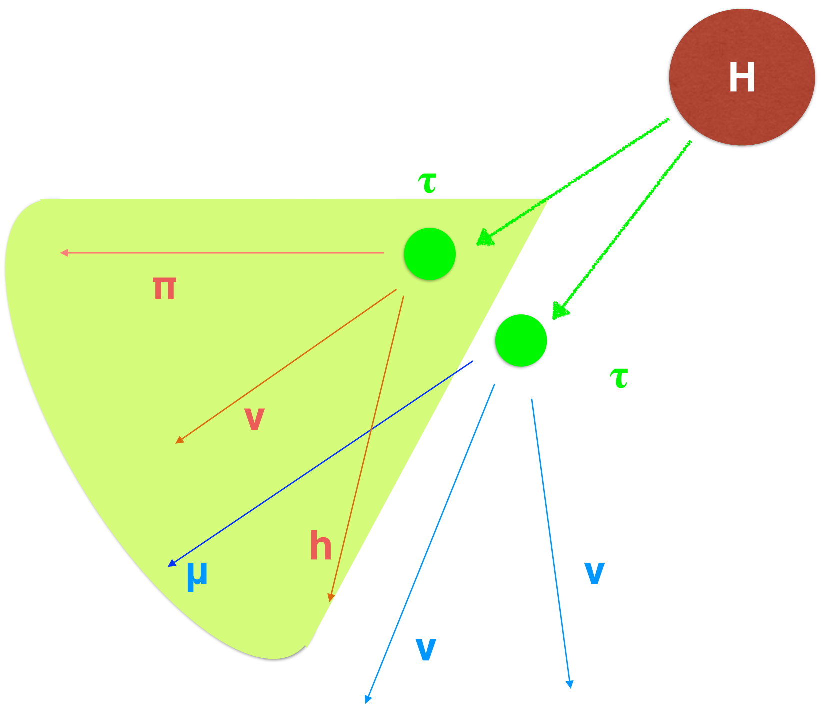


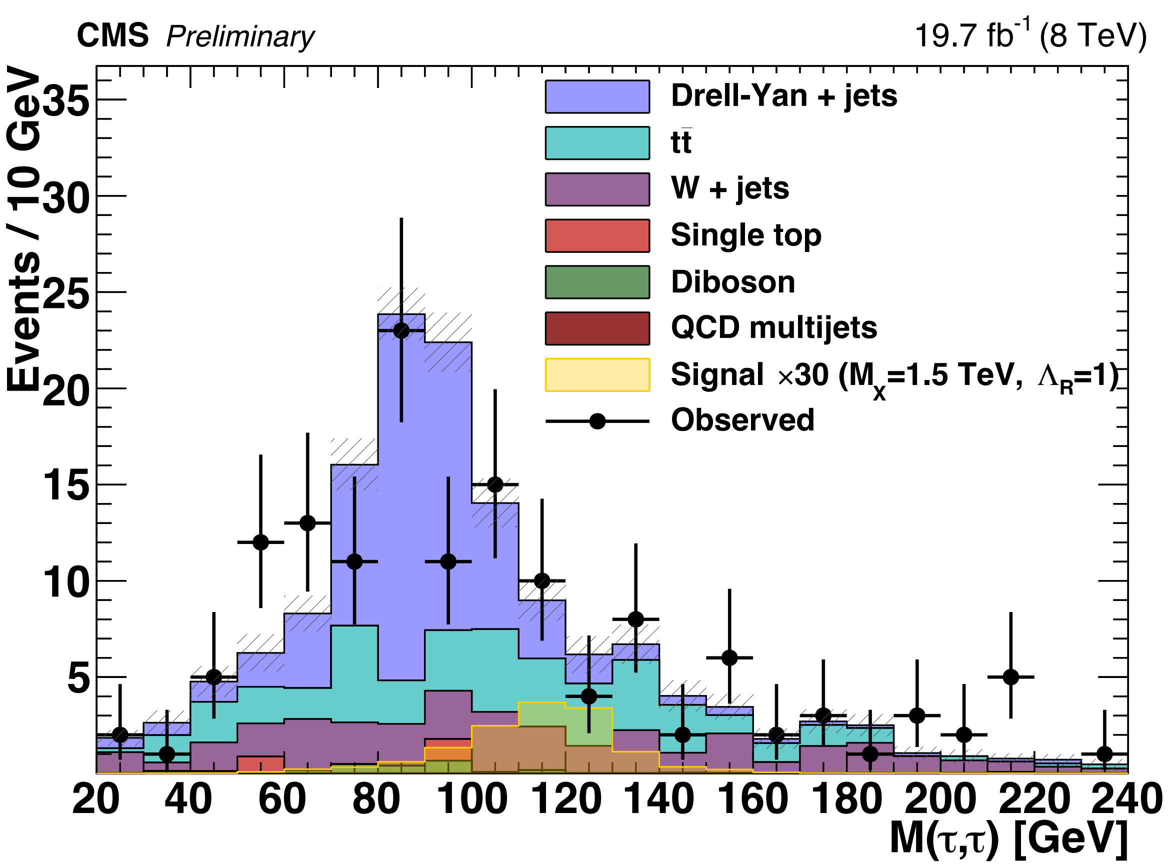


Higgs➜ττ tagging

- τ-lepton can decay hadronically and leptonically
- need to take into account **potential overlap** between the two **τ-leptons**
  - remove tracks/particles entering other isolation cone
- **discrimination against q-/g-jets**: MVA-based isolation
  - sum reconstructed particle energies in various cones around τ decay products
- **neutrinos** in decay cannot be reconstructed ➜ missing energy
  - ττ-reconstruction using templates from Monte Carlo simulation (SVfit)

currently only 8 TeV results

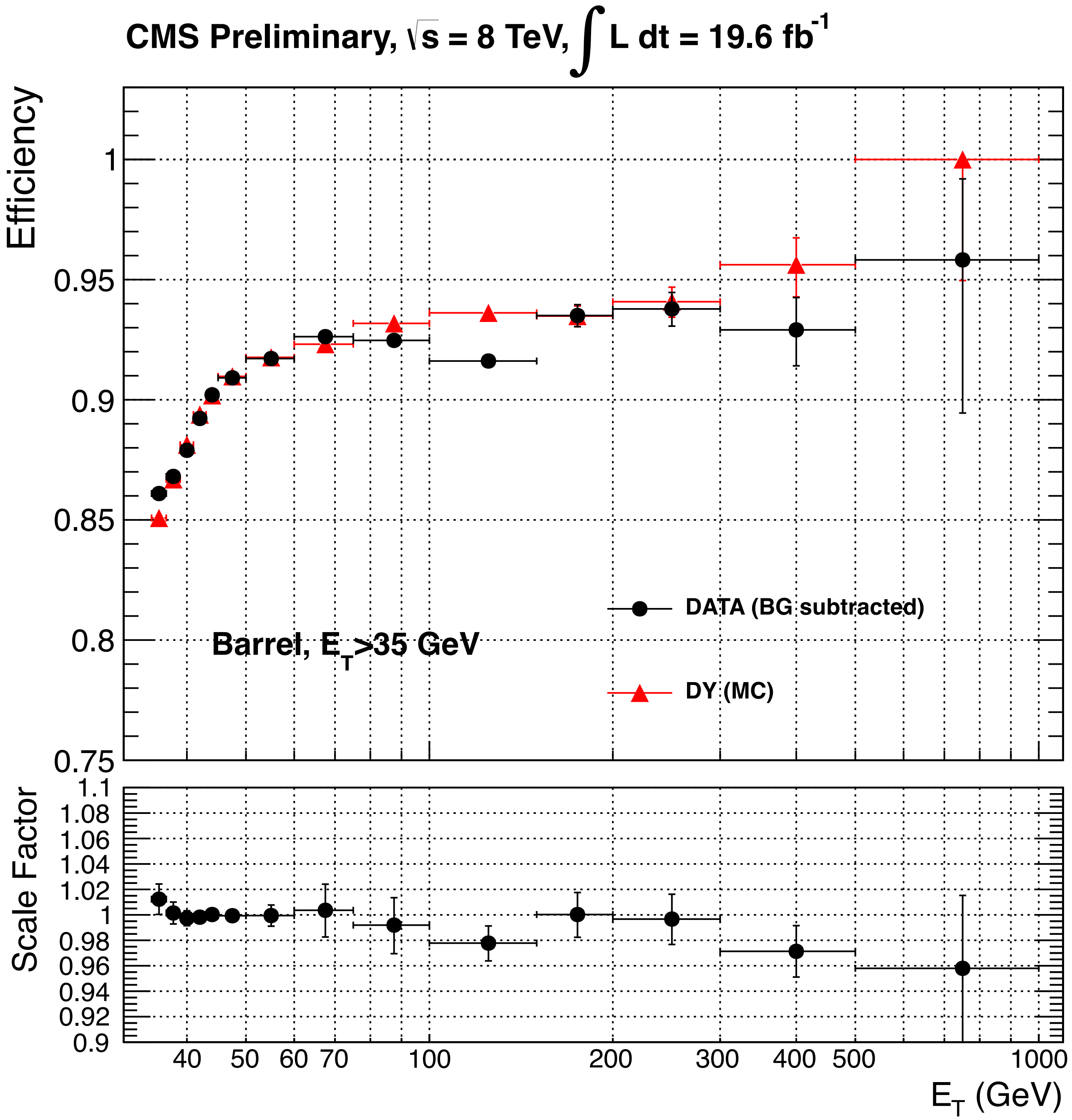




Lepton (muon + electron) reconstruction

- two **isolation** issues:
  - radiation from highly energetic leptons spoils isolation
  - leptons spoil each other's isolation
- employ **dedicated high-p_T algorithms** to preserve high efficiency
- loosen selection criteria one of the leptons in Z➜ll decays
- leptonic W-decay: need to recover z-component of neutrino
  - use W-mass constraint for reconstruction

High energy electron ID (HEEP)



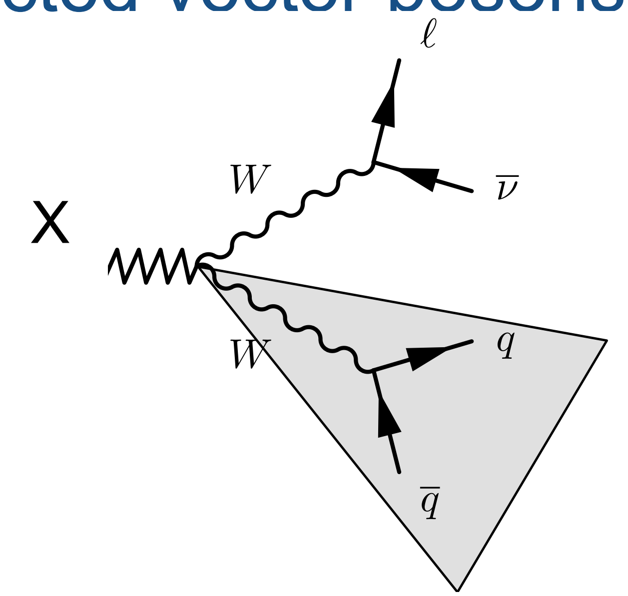
Strategy/event selection for VV analyses

	VV ➜ qqqq analysis	VW ➜ qqlν analysis	VZ ➜ qqll analysis
trigger	H_T trigger (800 GeV) or jet+groomed mass	single lepton trigger (e/μ p_T > 105/45 GeV)	
lepton(s)	—	HEEP e/high-p_T μ	special iso/ID for 2nd lepton
V-jet	anti-k_T R=0.8, p_T > 200 GeV, exploit substructure τ_2/τ_1, use groomed mass		
V boson candidate(s)	$\Delta\eta < 1.3$	reconstruct leptonic V, p_T > 200 GeV use mass window	
X	additional cuts on separation of bosons in $\Delta\phi$ and ΔR reconstruct X using both reconstructed vector bosons		

search for bump in m_{VV} distribution

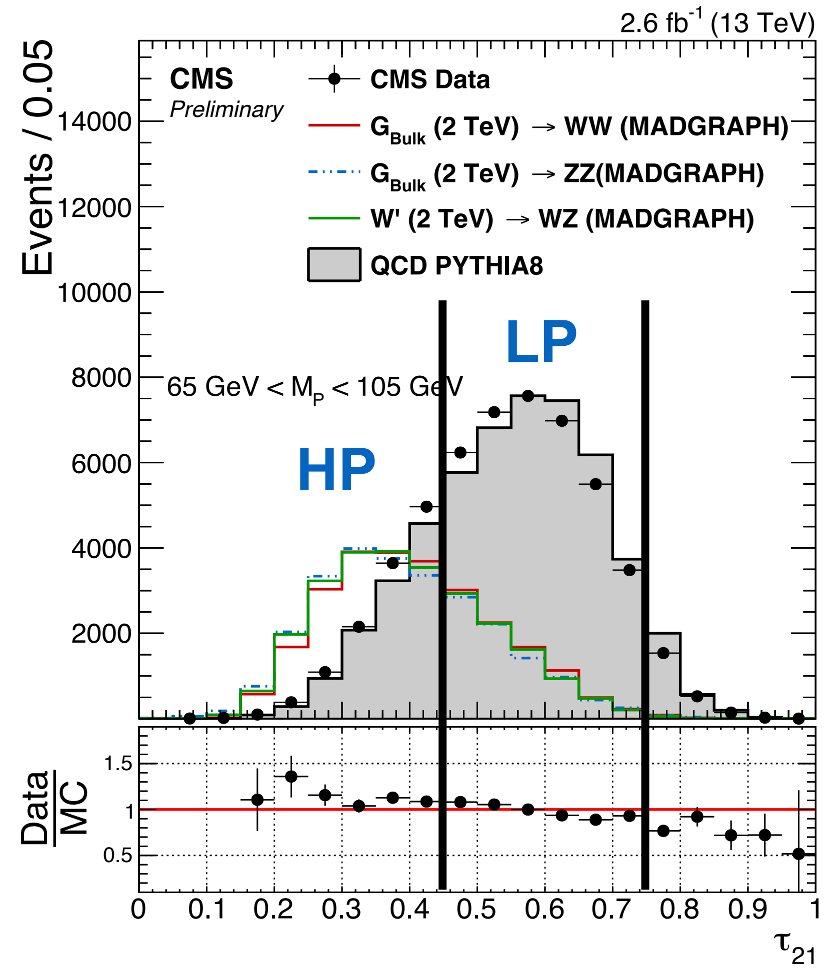
diboson-like topology:

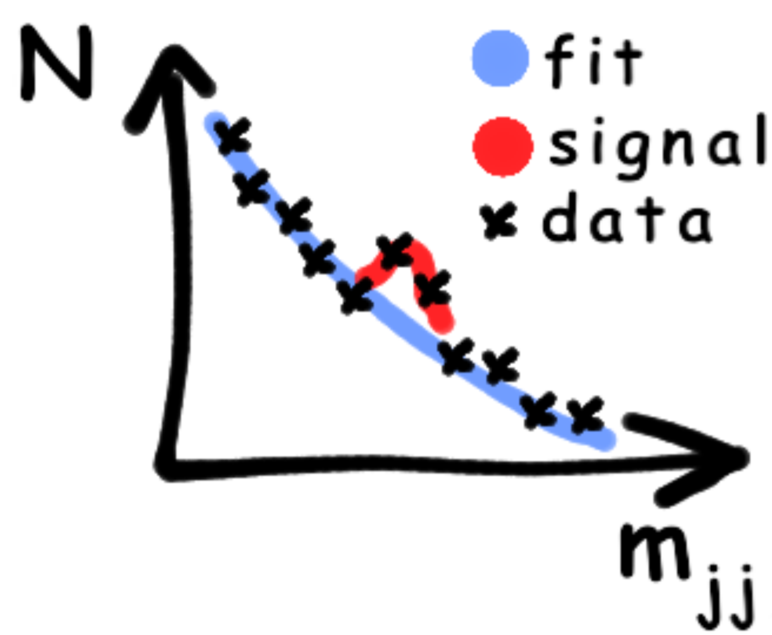
- $\Delta R(\ell, W_{had}) > \pi/2$
- $\Delta\phi(W_{had}, E_T^{miss}) > 2$
- $\Delta\phi(W_{had}, W_{lep}) > 2$



Dijet (VV) analysis

- **trigger** at ~100% efficiency at **m_{JJ} > 1 TeV**
  - apply cut on reconstructed dijet system
- define different **τ_2/τ_1 regions**:
  - high purity to suppress background
  - low purity to recover signal efficiency at high masses
- **split W and Z** samples based on pruned jet mass (65-85, 85-105 GeV)
- still dominated by **QCD multi-jet** events
- difficult to obtain sufficient MC simulation statistics
- need a data-driven approach
- exponentially falling spectrum: use **fit function**





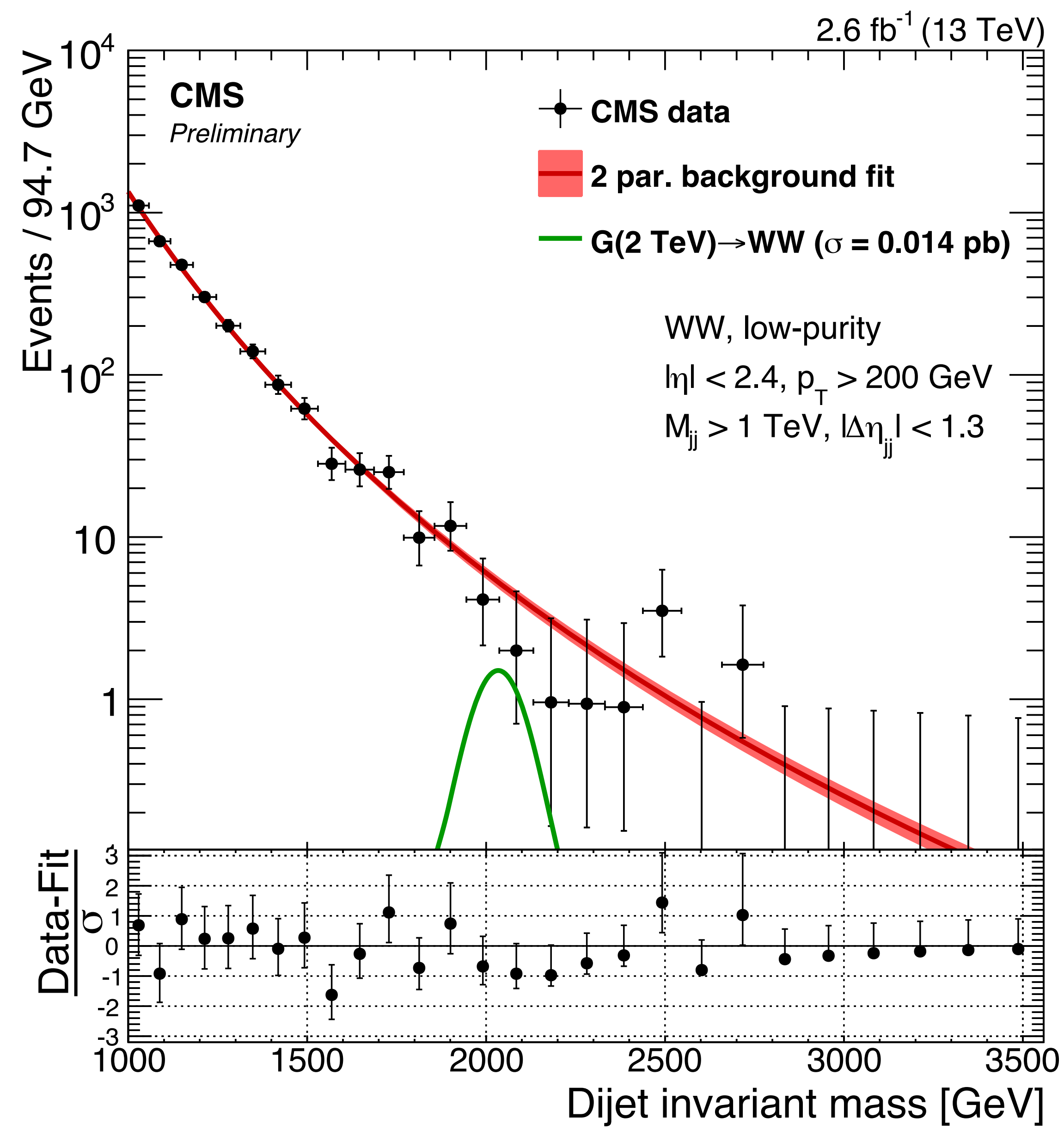
VV background estimation

- naively, fitting the m_{JJ} spectrum could swallow signal
- also, need to **avoid claiming false discovery** (in particular in tail)
- fit function:

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

2 parameters — 3 parameters

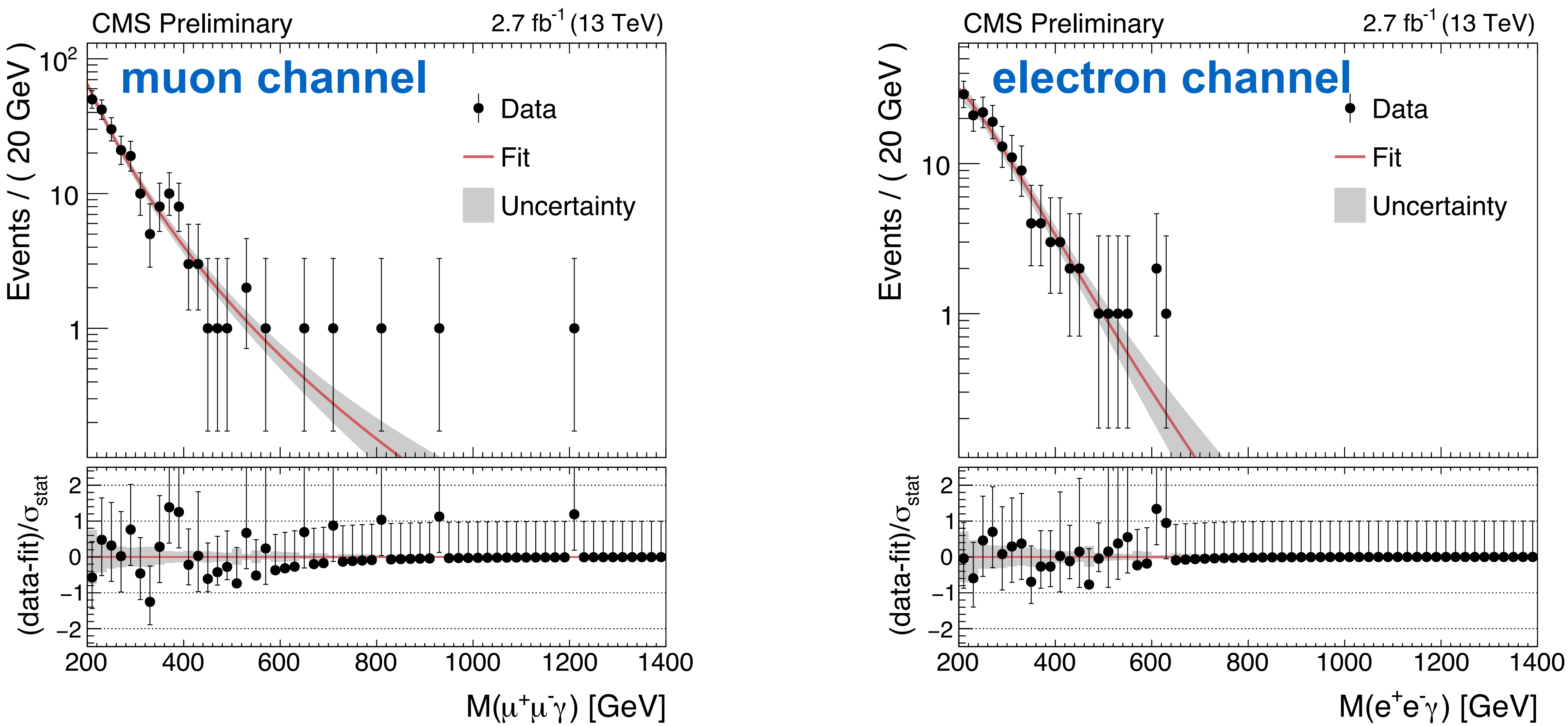
- number of free parameters determined by **F-test**:
  - check if quality of fit improves by > 10% confidence level
  - if not, stick with current fit function
- **extensive bias tests** conducted
- combined **signal+background fit** performed



very similar background estimation strategy as for diphoton resonance search

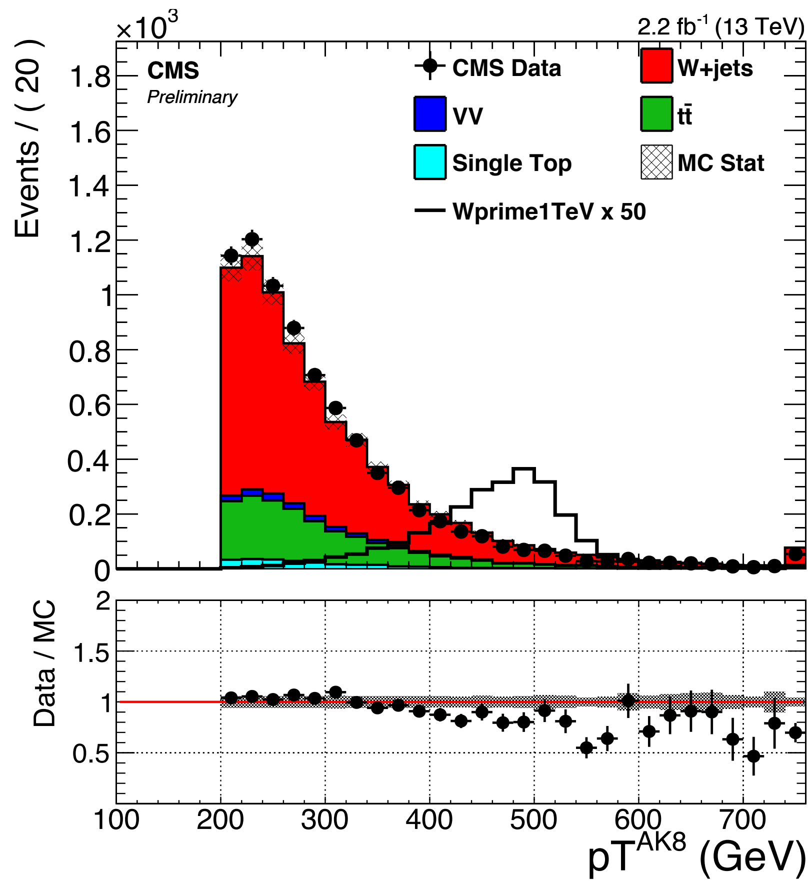
Intermezzo: Zɣ search overview

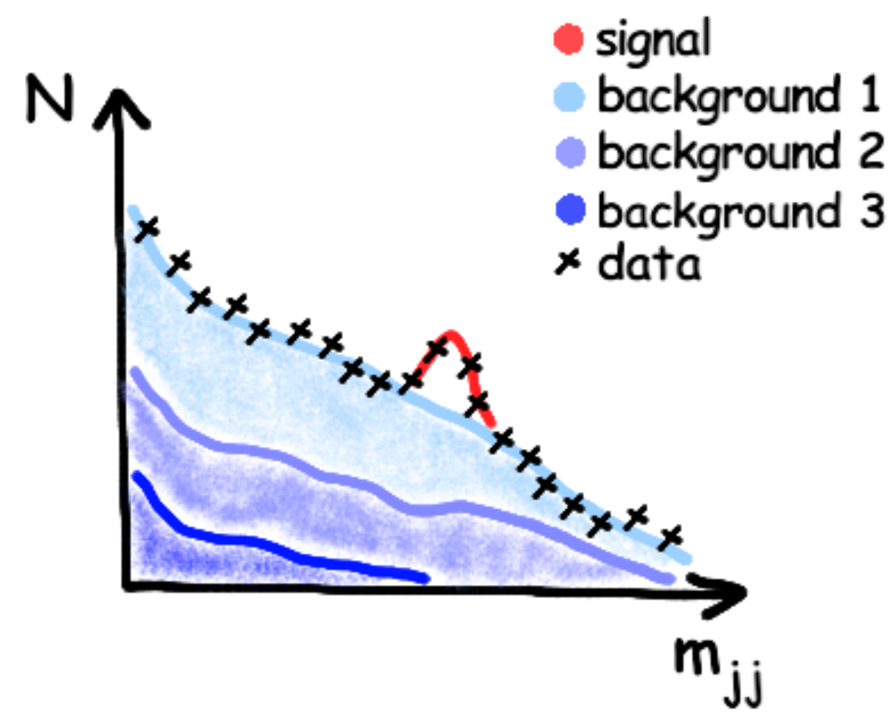
- recently published Z➜ll + ɣ search
- inspired by ɣɣ „excess“
- same photon ID as ɣɣ search, dilepton ID as in ZV search, fit background
- limited by statistics, **no significant excess**



VW/VZ analysis overview

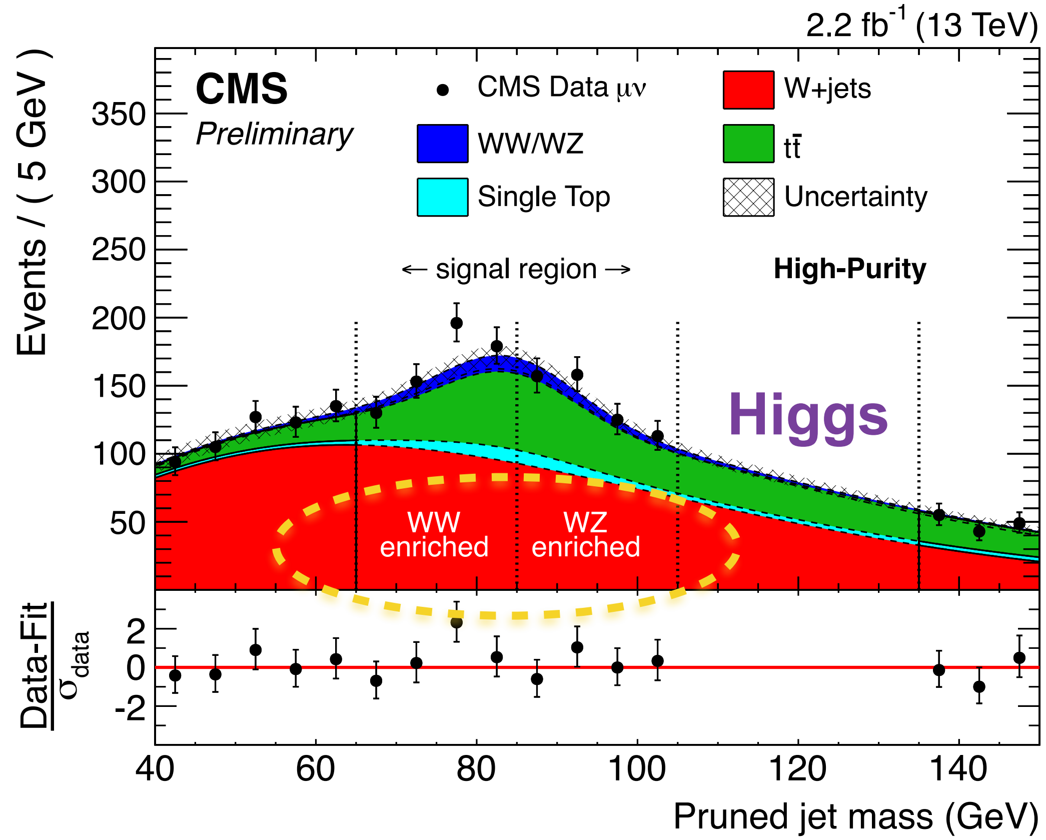
- for 2015 data, two separate analyses performed:
  - „low mass“: 600-1000 GeV
  - „high mass“: 1-4 TeV
  - VZ analysis not yet public
- difference **low** vs. **high mass**:
  - lower boost ➜ can use isolated lepton triggers with 27 GeV thresholds
- requiring an isolated lepton suppresses QCD multi-jet background significantly
- dominant backgrounds:
  - Drell-Yan/W+Jets
  - top-antitop quark production
- can **estimate individual background components** from sidebands

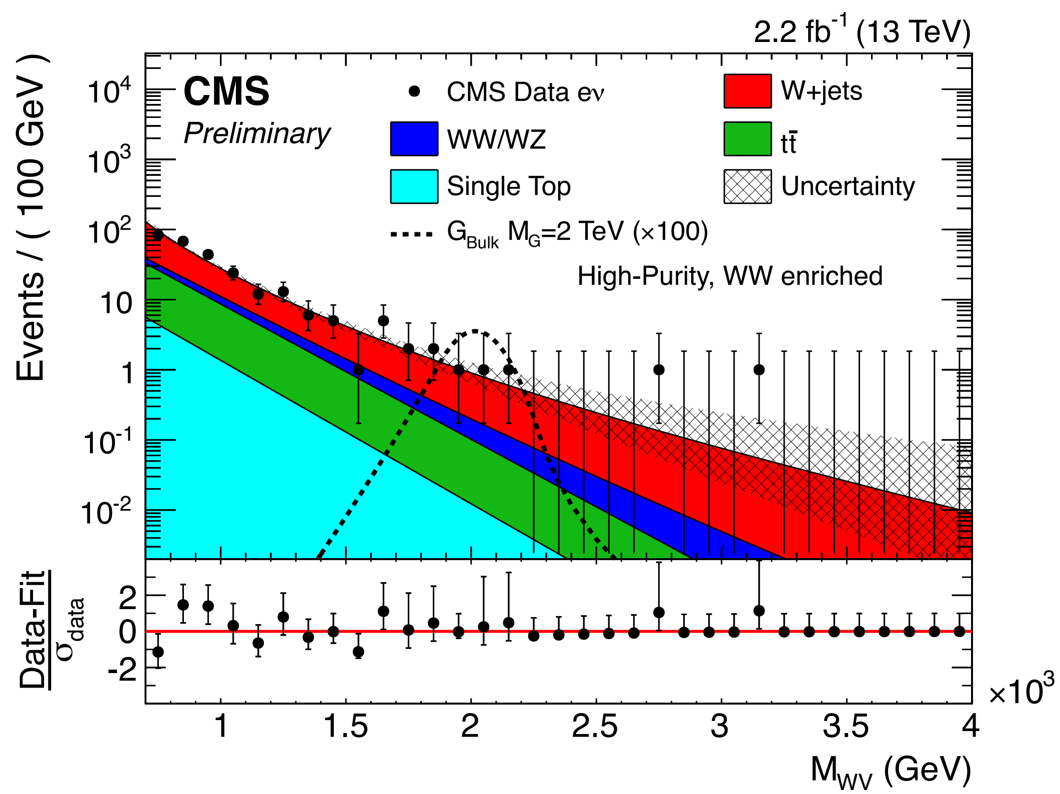




VW analysis background estimation

- statistics in MC simulated samples still limited
- furthermore, analysis performed in **extreme phase space**
- use **pruned mass sidebands** (40-65 GeV, 135-160 GeV) to exploit correlation between pruned 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

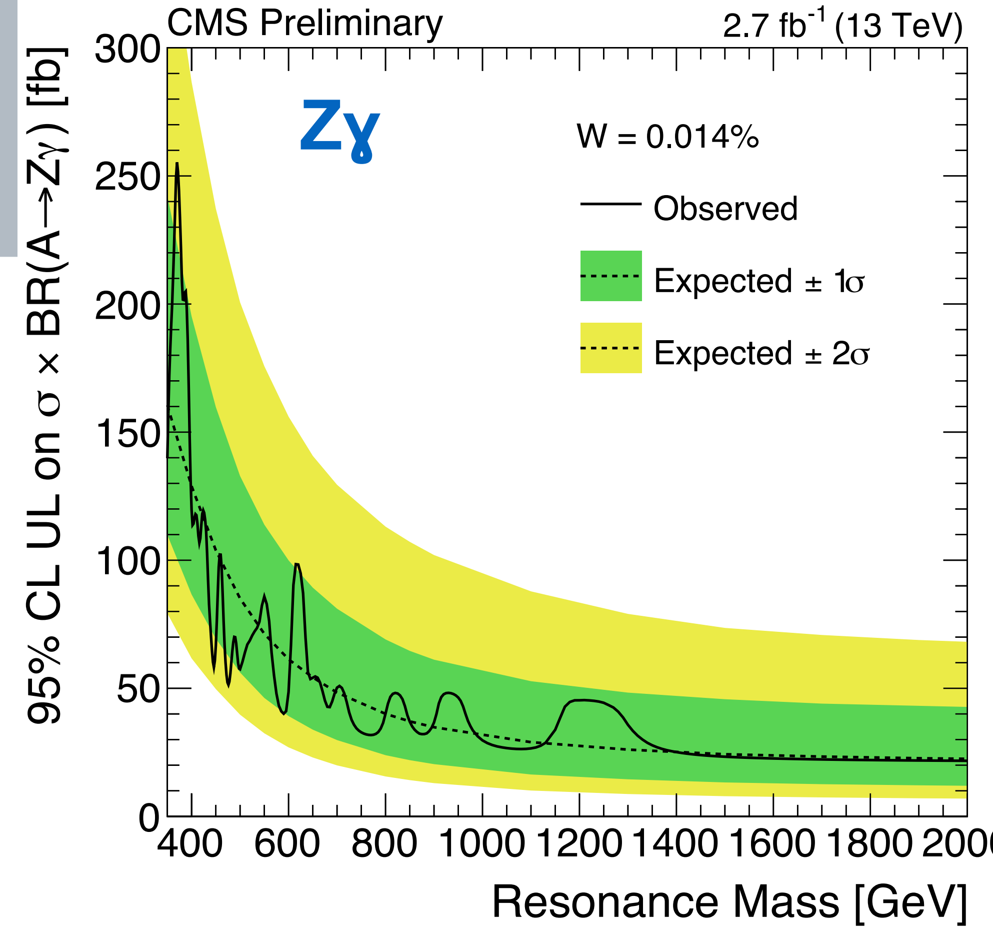


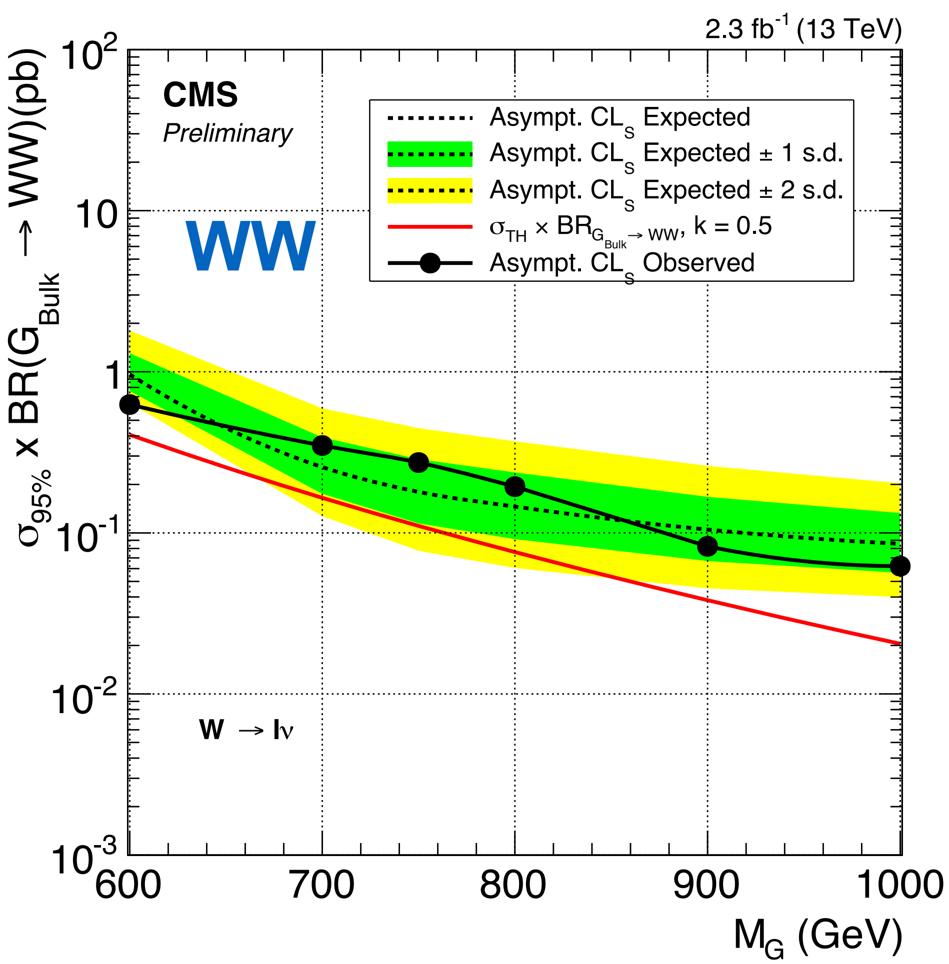


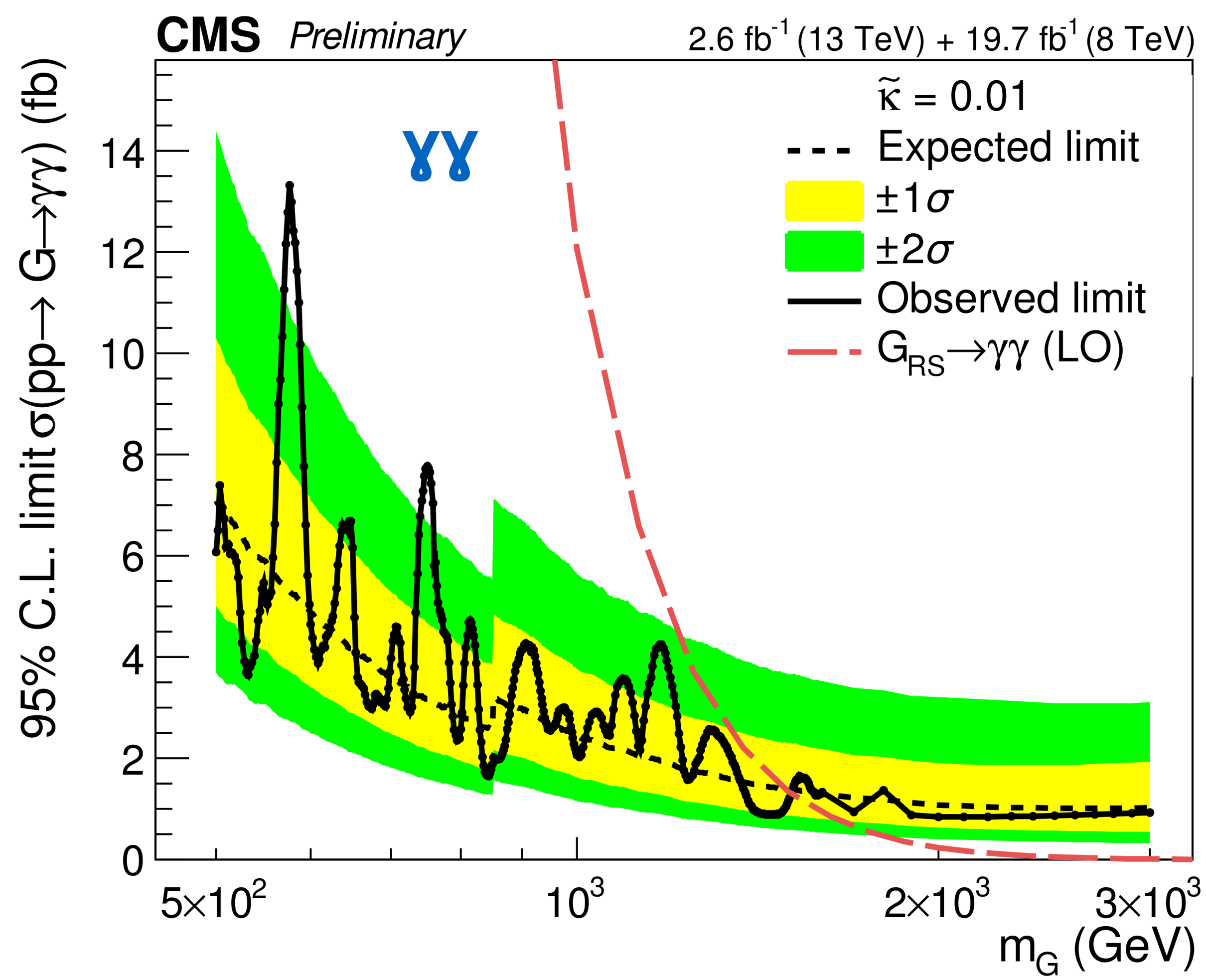
Intermezzo: γγ vs. Zγ vs. WW

- limits for **narrow resonances**
  - caveat: slightly different models used
- minimal upward fluctuations?

CMS-PAS-EXO-15-004
CMS-PAS-EXO-16-019
CMS-PAS-B2G-16-004

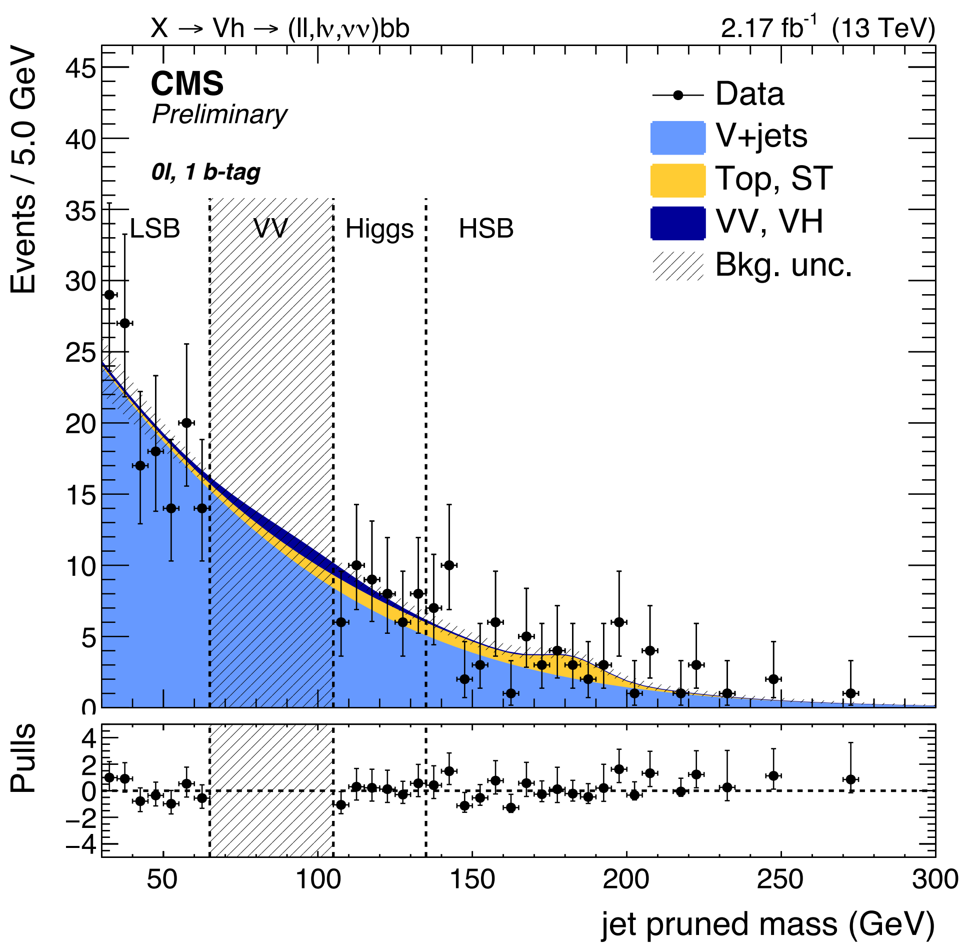




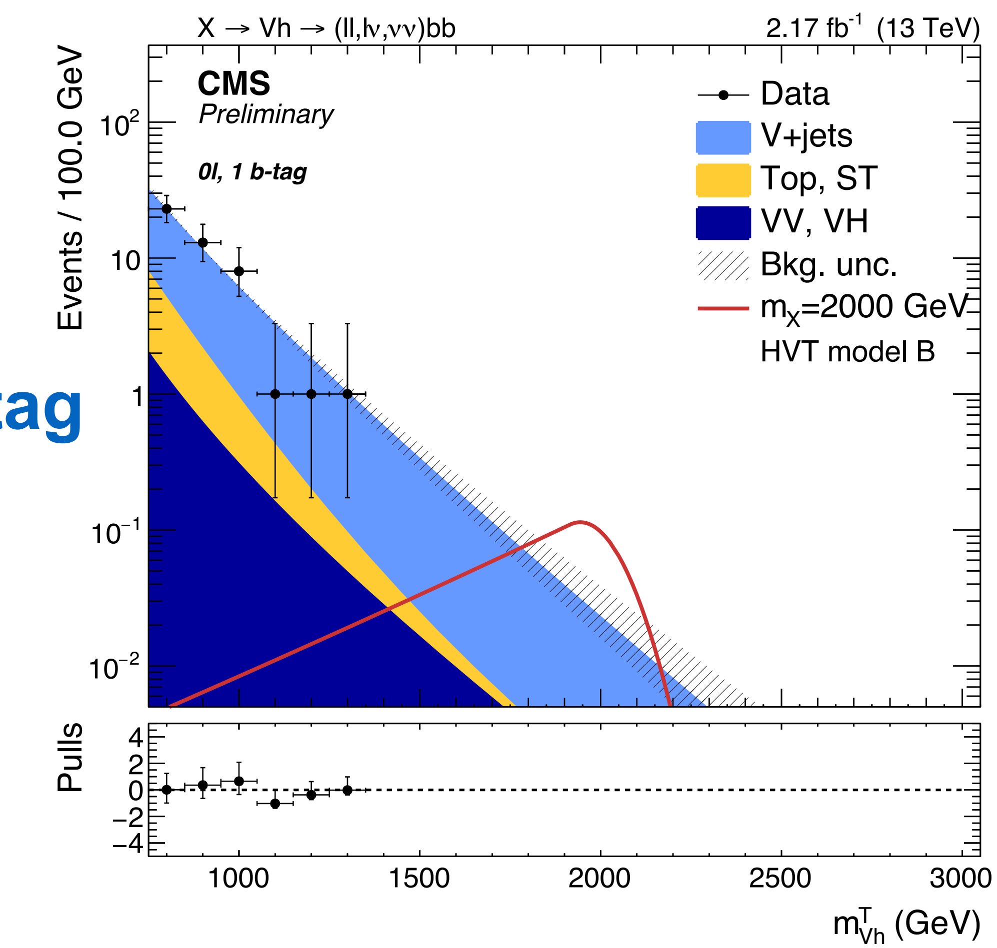


VH analysis overview

- 2015 data: H➔bb, leptonic W/Z decays
  - W➔lν
  - Z➔ll
  - **Z➔νν**
- categorise in **single** and **double subjet b-tag** categories
- same background estimation method as for VW search

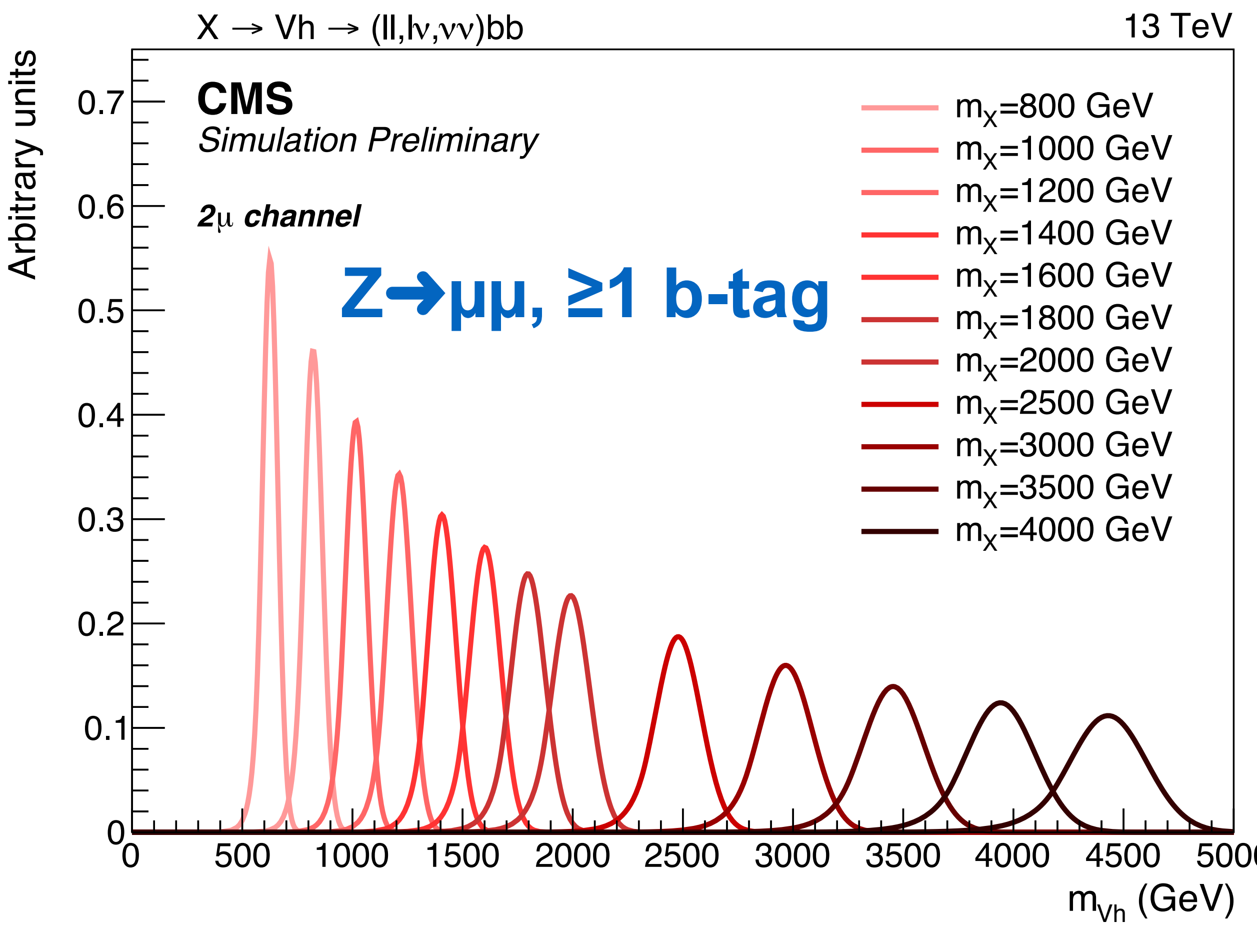


Z➔νν, 1 b-tag



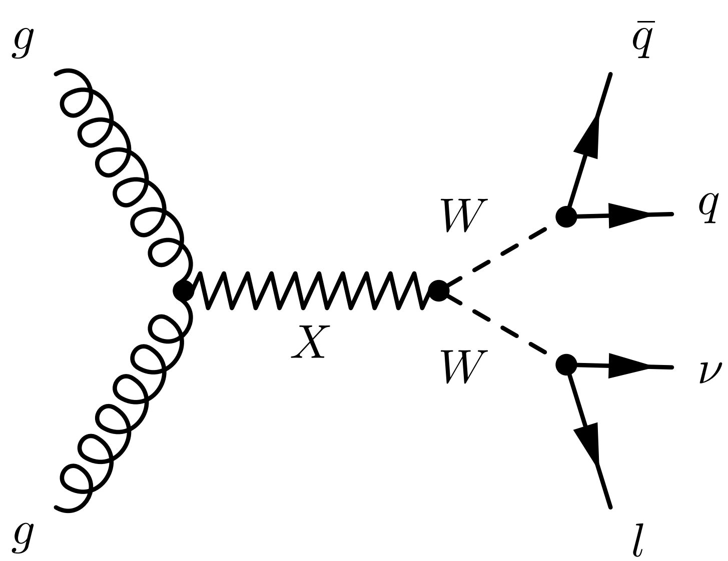
Signal modelling and uncertainties

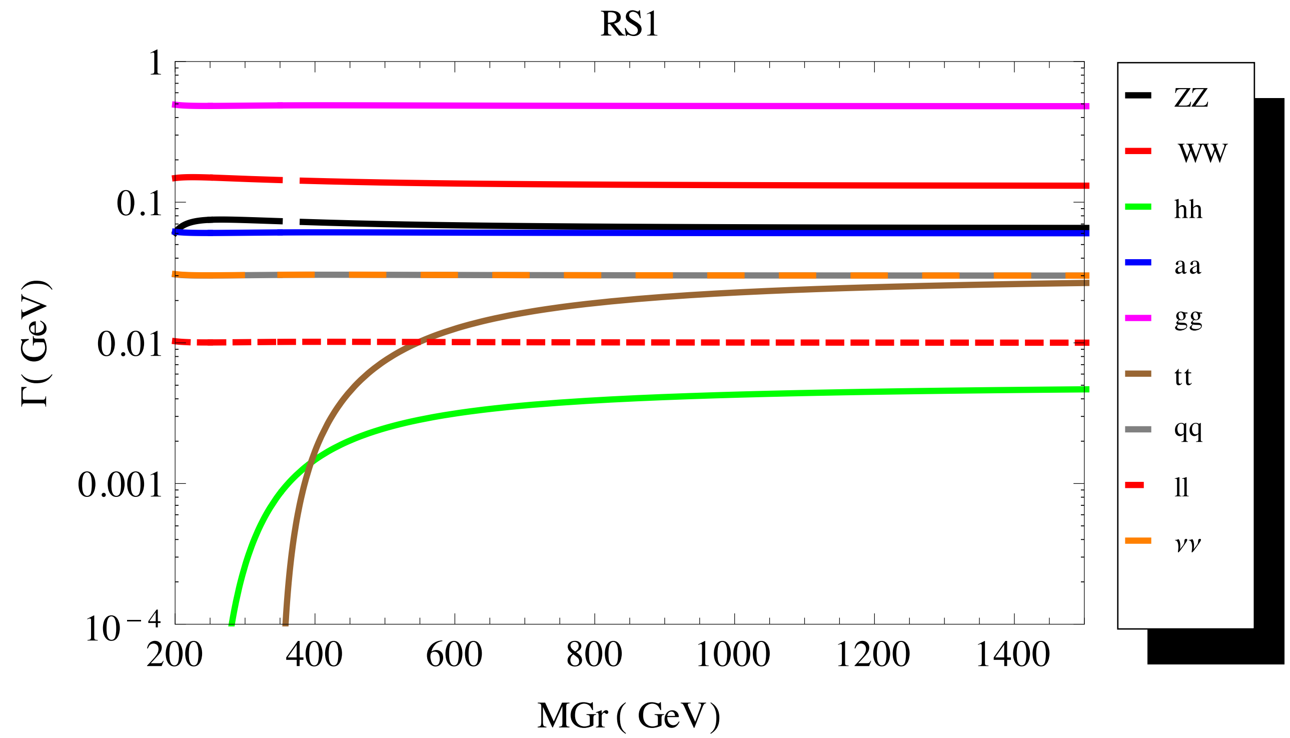
- depending on spin of new particles, **polarisation of bosons** different
- Bulk graviton (spin-2) and W'/Z' (spin-1) models primarily couple to **longitudinal components of W/Z**
- analytical description of **signal shapes** based on fully simulated benchmark mass points
  - double-sided Crystal-Ball function
  - linearly interpolation between benchmark points
- signal efficiency up to 15% depending on analysis category
- largest uncertainties:
  - background estimation
  - jet energy and mass scale
  - boson-tagging

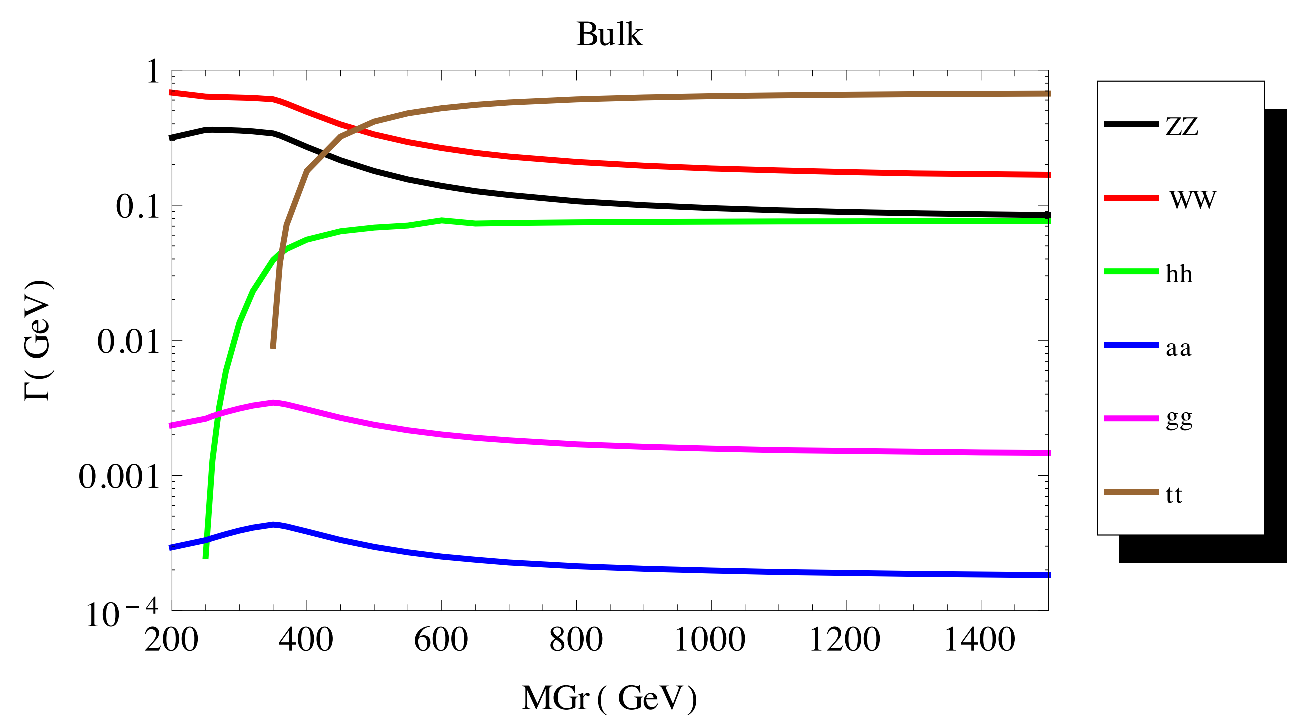


Model interpretation

- several different analyses performed
- advantageous to use **common benchmark models** for easier interpretation
- need to know **individual couplings to bosons**
  - individual analysis: e.g. σ(gg➔G➔WW)
  - combination: e.g. σ(gg➔G)
  - mind also production mechanism





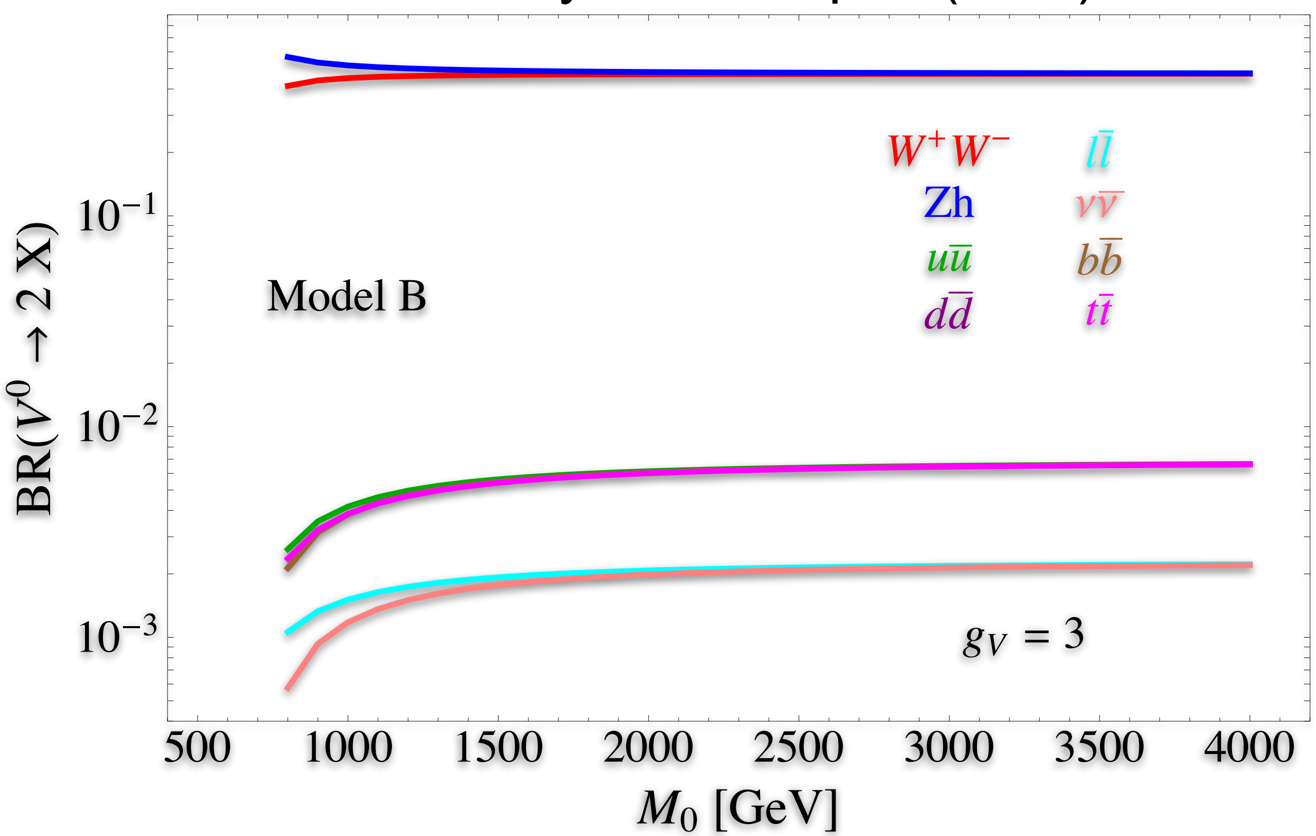


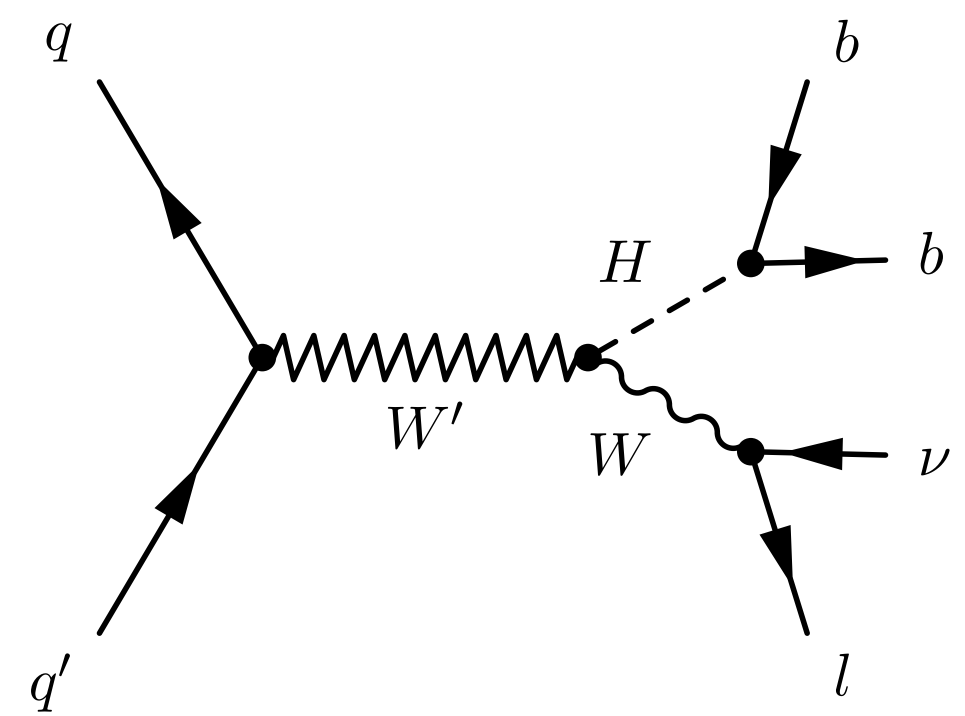
10.1007/JHEP09(2014)060

Model tuning

- models described before can be tuned by a **handful of parameters**
- bulk graviton:
  - mass of graviton
  - coupling constant determining production cross section and width
- heavy vector triplet (HVT):
  - phenomenological Lagrangian
  - describes production and decay of heavy spin-1 resonances
  - 4 parameters for resonance mass, interaction strength, couplings to bosons and fermions
  - focus here on „Model B“ with **enhanced couplings to bosons**

heavy vector triplet (HVT)





Putting it all together

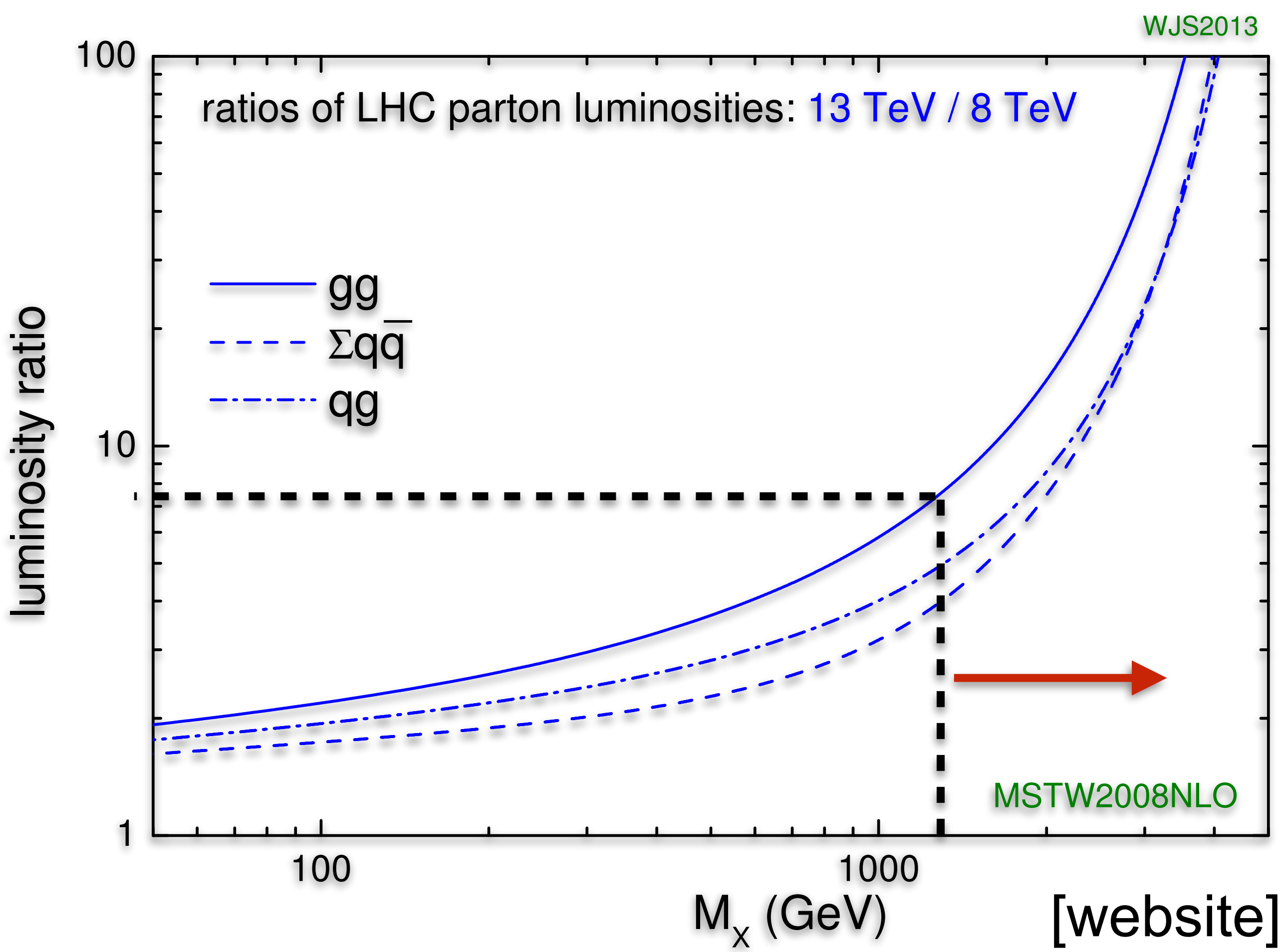
- currently, larger number of channels has been covered with **8 TeV data**
- **differentiate between final states** (number of leptons and jets)
- **violet** analyses interpreted in dark matter scenarios

	$W \to l\nu$	$Z \to ll$	$V \to qq$	$Z \to \nu\nu$	$H \to qqqq$	$H \to \tau\tau$	$H \to bb$	$H \to \gamma\gamma$
$W \to l\nu$			13 TeV				13 TeV	
$Z \to ll$			13 TeV				13 TeV	
$V \to qq$	13 TeV	13 TeV	13 TeV	13 TeV				
$Z \to \nu\nu$			13 TeV				13 TeV	
$H \to qqqq$								
$H \to \tau\tau$								
$H \to bb$	13 TeV	13 TeV		13 TeV				
$H \to \gamma\gamma$								

- several analyses repeated and improved w.r.t. 8 TeV (**considering m_X > 1TeV**)
- several more to come this year

8 vs. 13 TeV

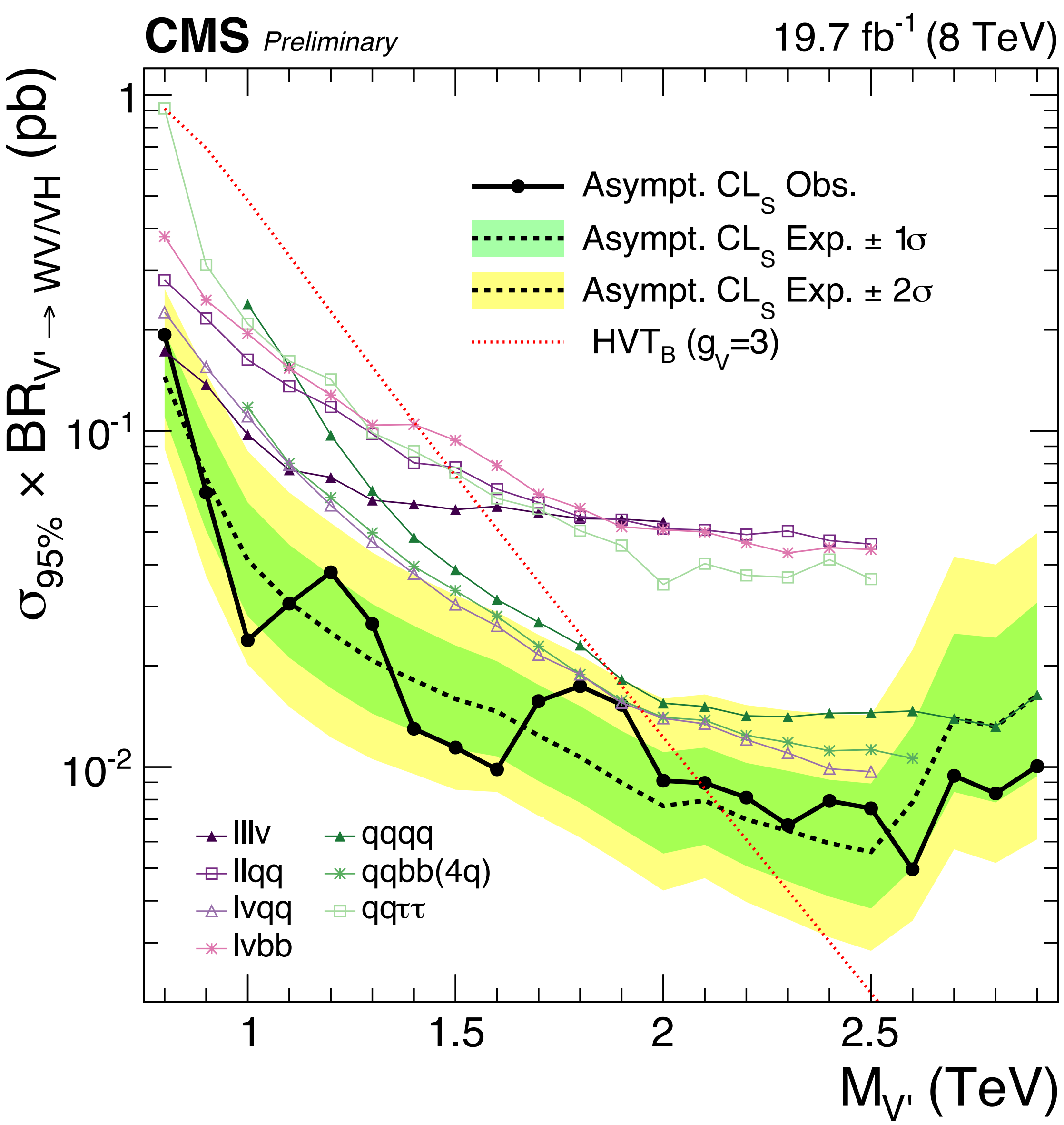
- While 13 TeV has opened up new energy regime, integrated luminosity recorded in 2015 significantly below the one of 2012
  - 20 fb^{-1} vs. ~3 fb^{-1}
- LHC is **hadron collider** - $\sqrt{s}$ ≠ energy available in collision
- need to consider **parton luminosities**
- exceed 2012 reach with 2015 data already at 1-2 TeV resonance mass
- nevertheless, worthwhile **combining results**

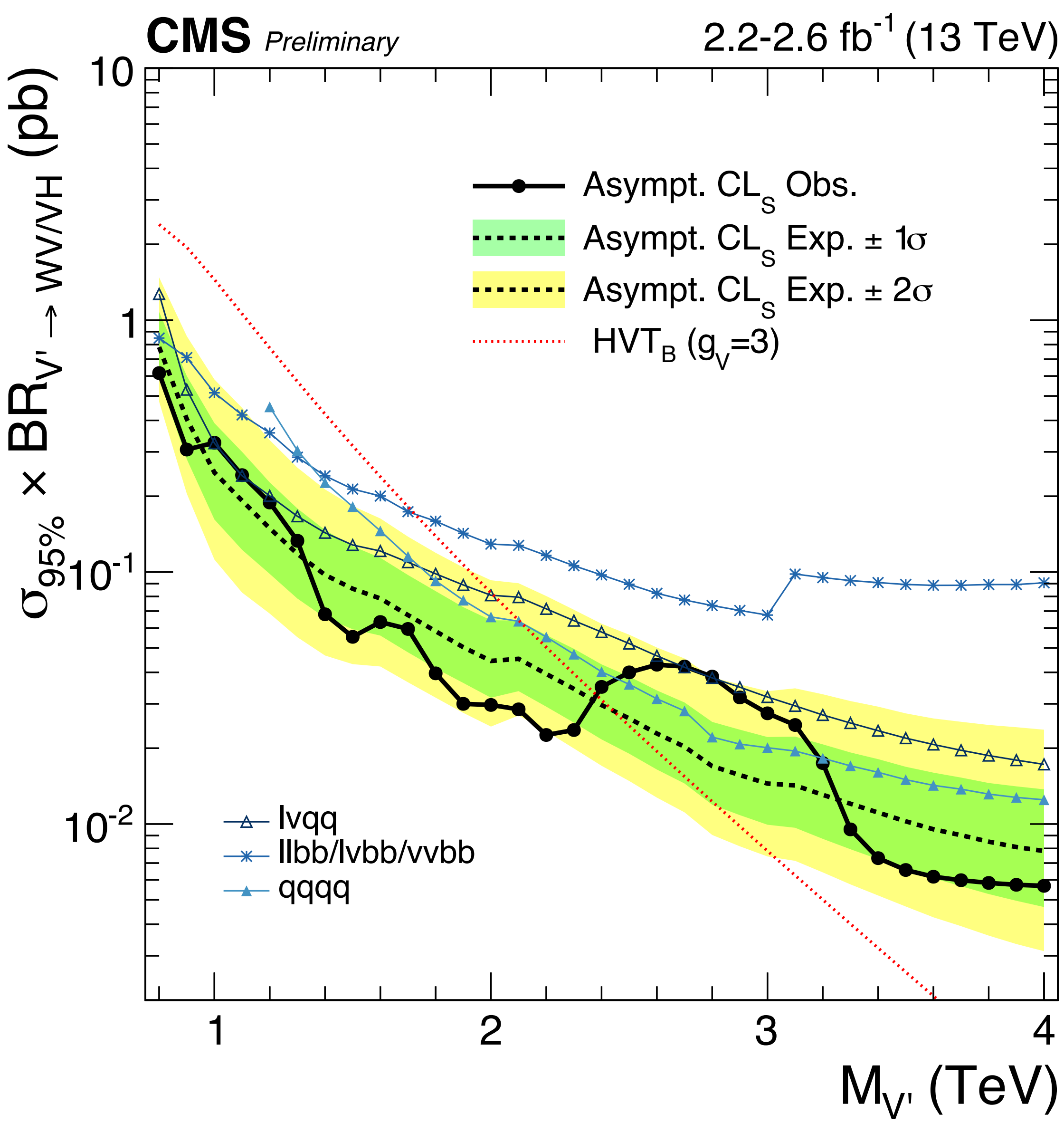


[website]

Combination of diboson analyses

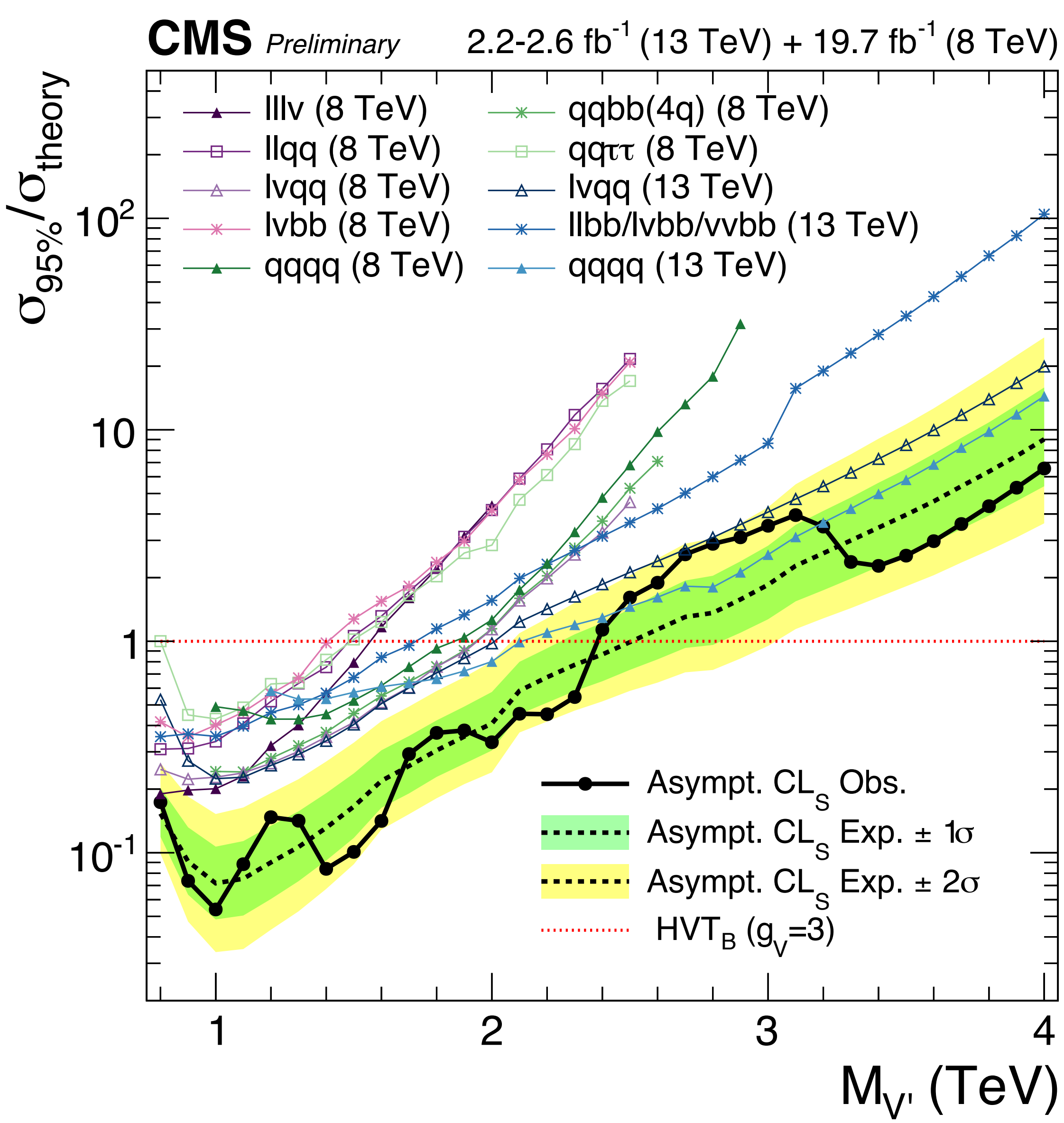
- example here: V' combination in HVT model B
- seven 8 TeV analyses, three at 13 TeV
- how to combine **upper cross section limits** from two different √s?

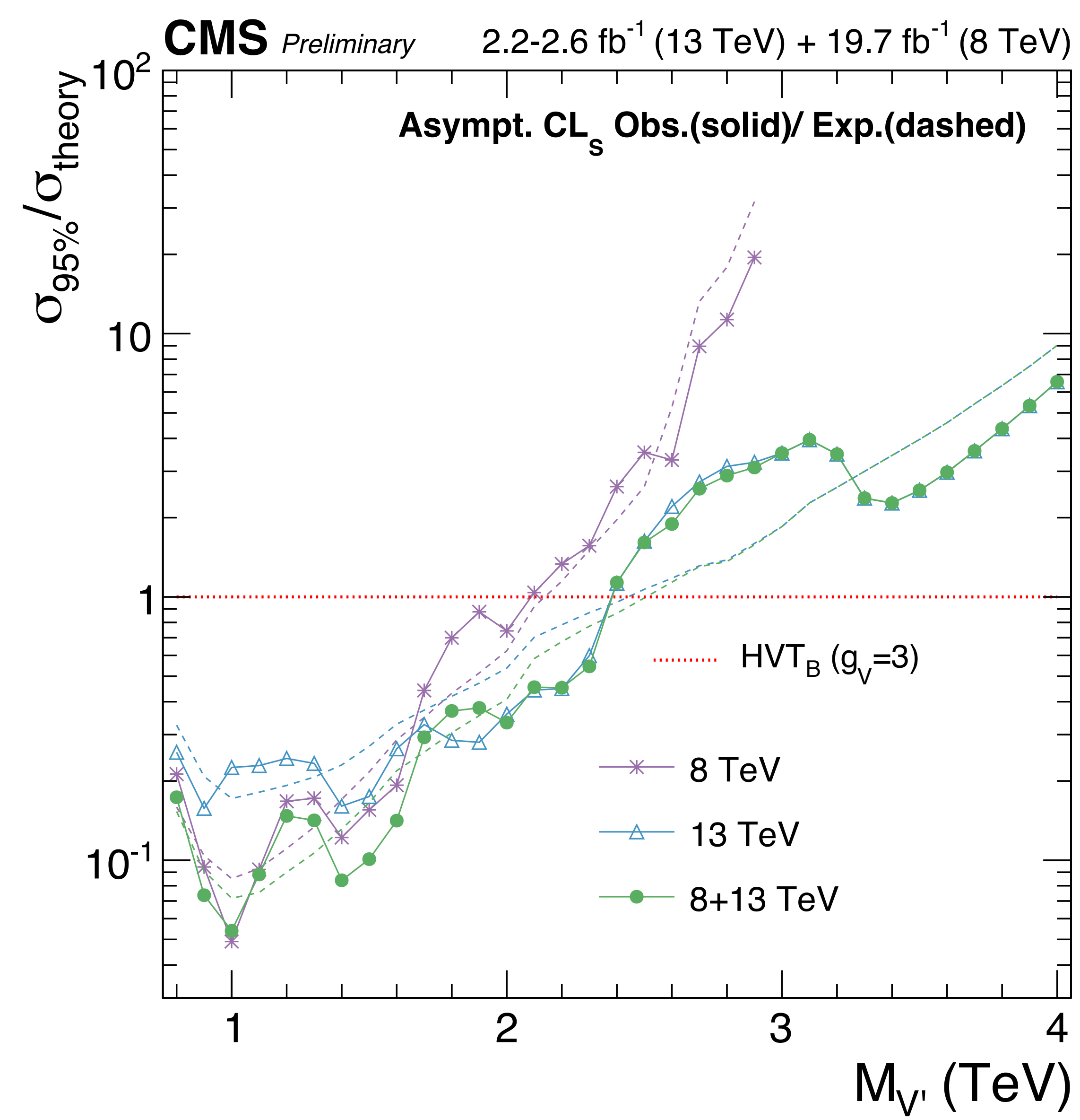




Combination of diboson analyses

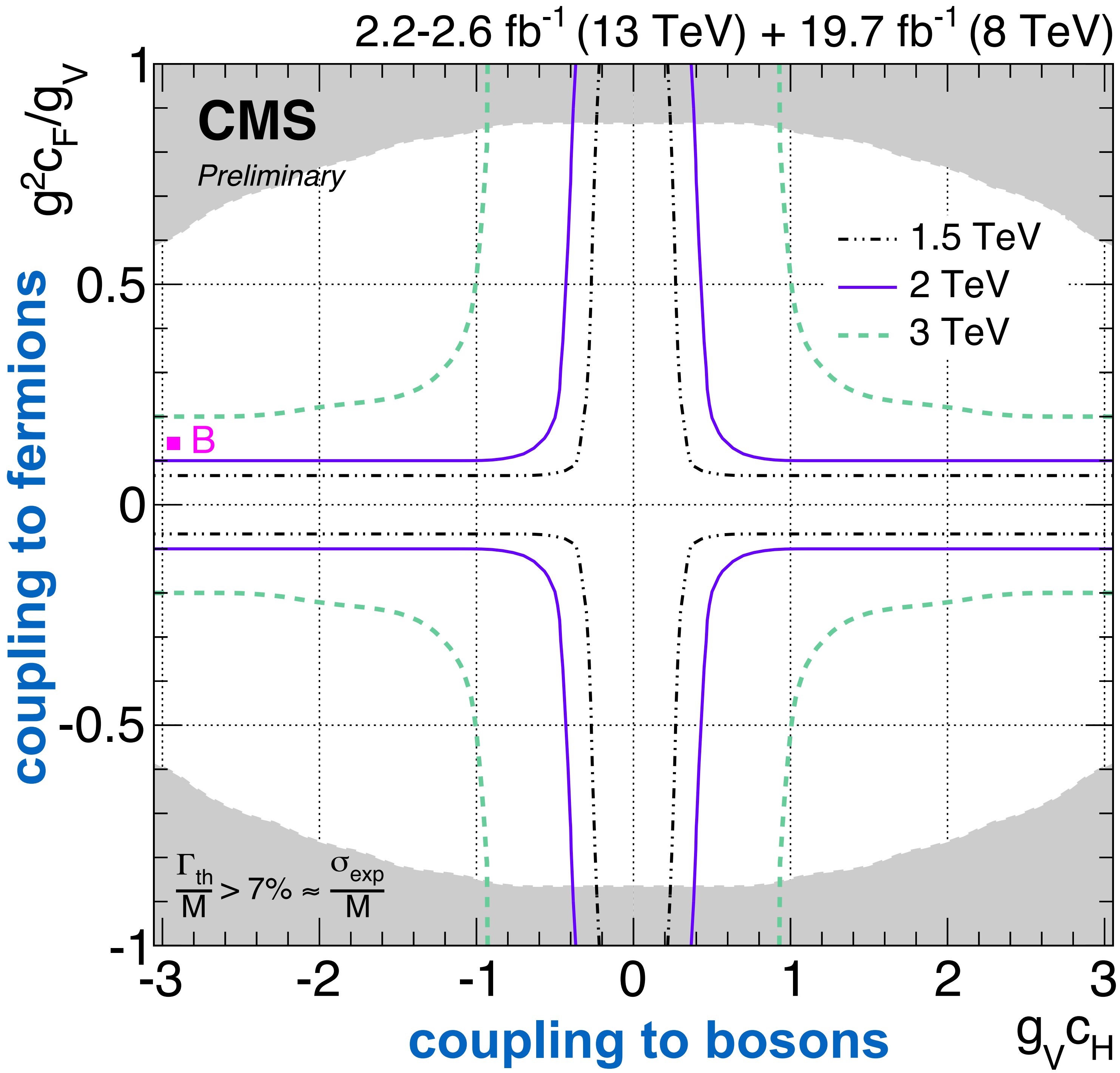
- convert cross section limits into **signal strength limits**
- 8+13 TeV limits comparable at lower masses, 13 TeV dominates high mass
- lower masses: leptonic analyses, higher masses: hadronic final states





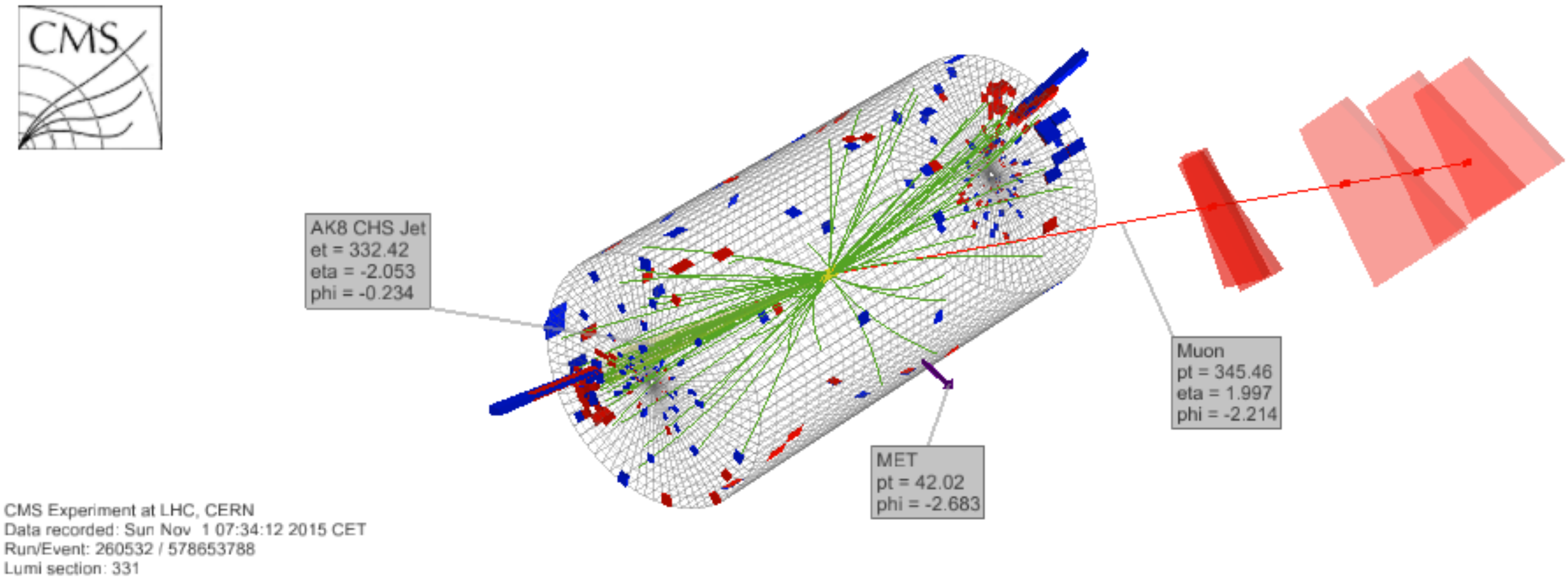
Combination of diboson analyses

- can translate observed limits into **exclusion contours** in the HVT couplings space
- additional input to theorists for model building



Summary

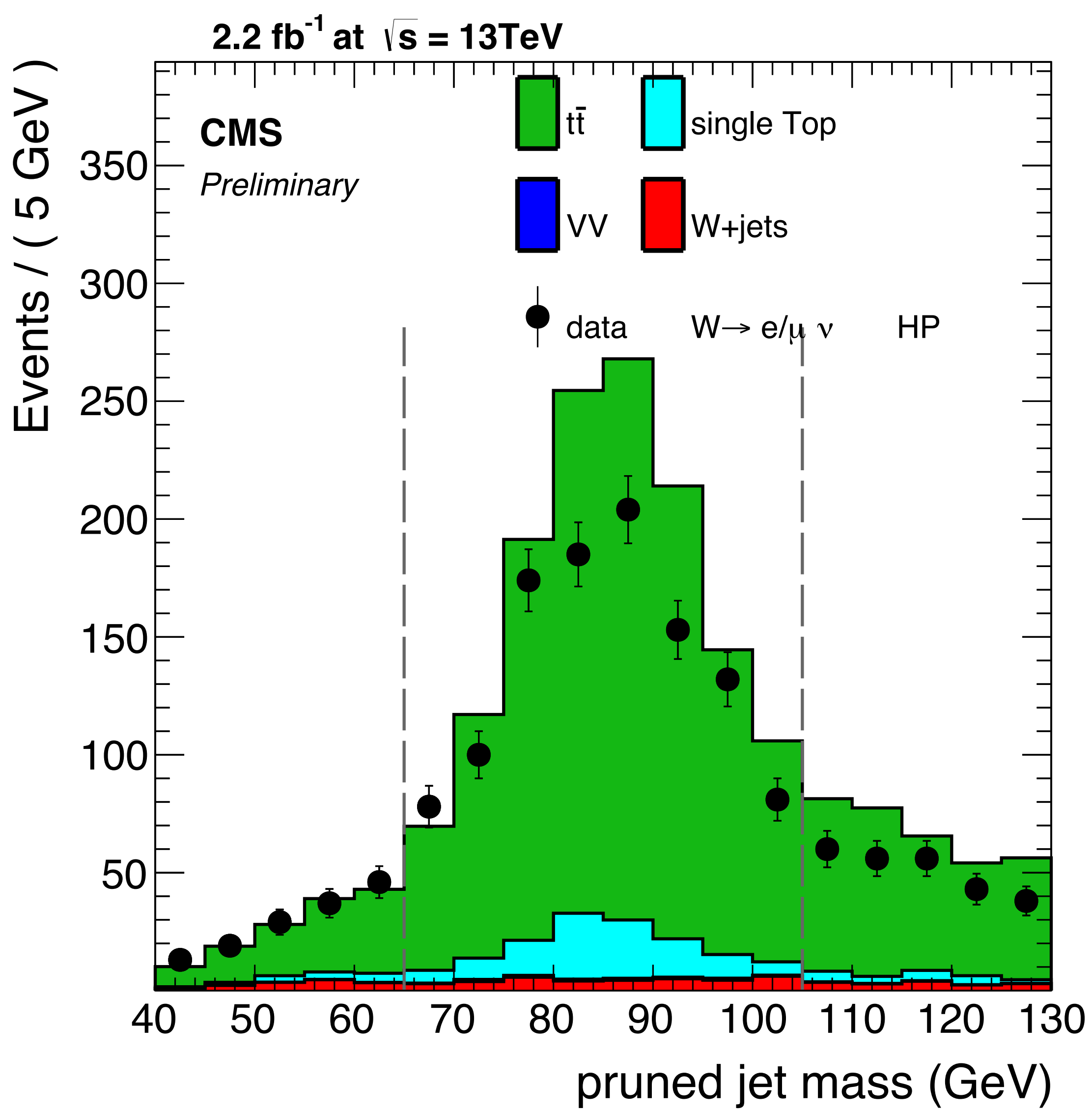
- **discovery** of a diboson resonance might solve hierarchy problem
- however, currently **no sign of new physics**
- 13 TeV results already exceed 8 TeV ones
- expect another **boost with 2016 data**

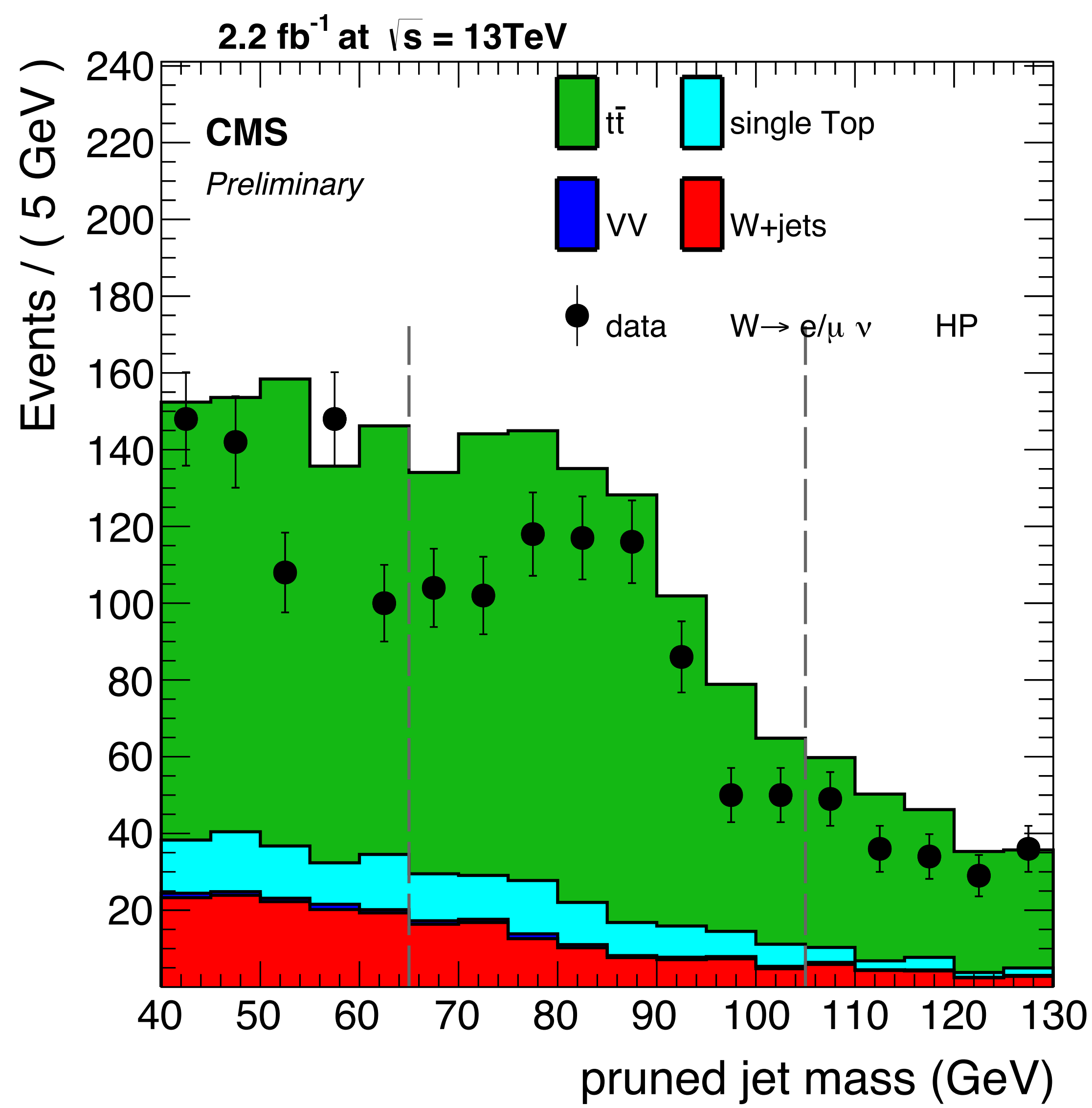


CERN

W-tagging calibration

- cutting on τ_2/τ_1-ratio ➜ need to know efficiency of cut in data and simulation
- select at generator level clean W-events and those that do not match
- perform simultaneous fit





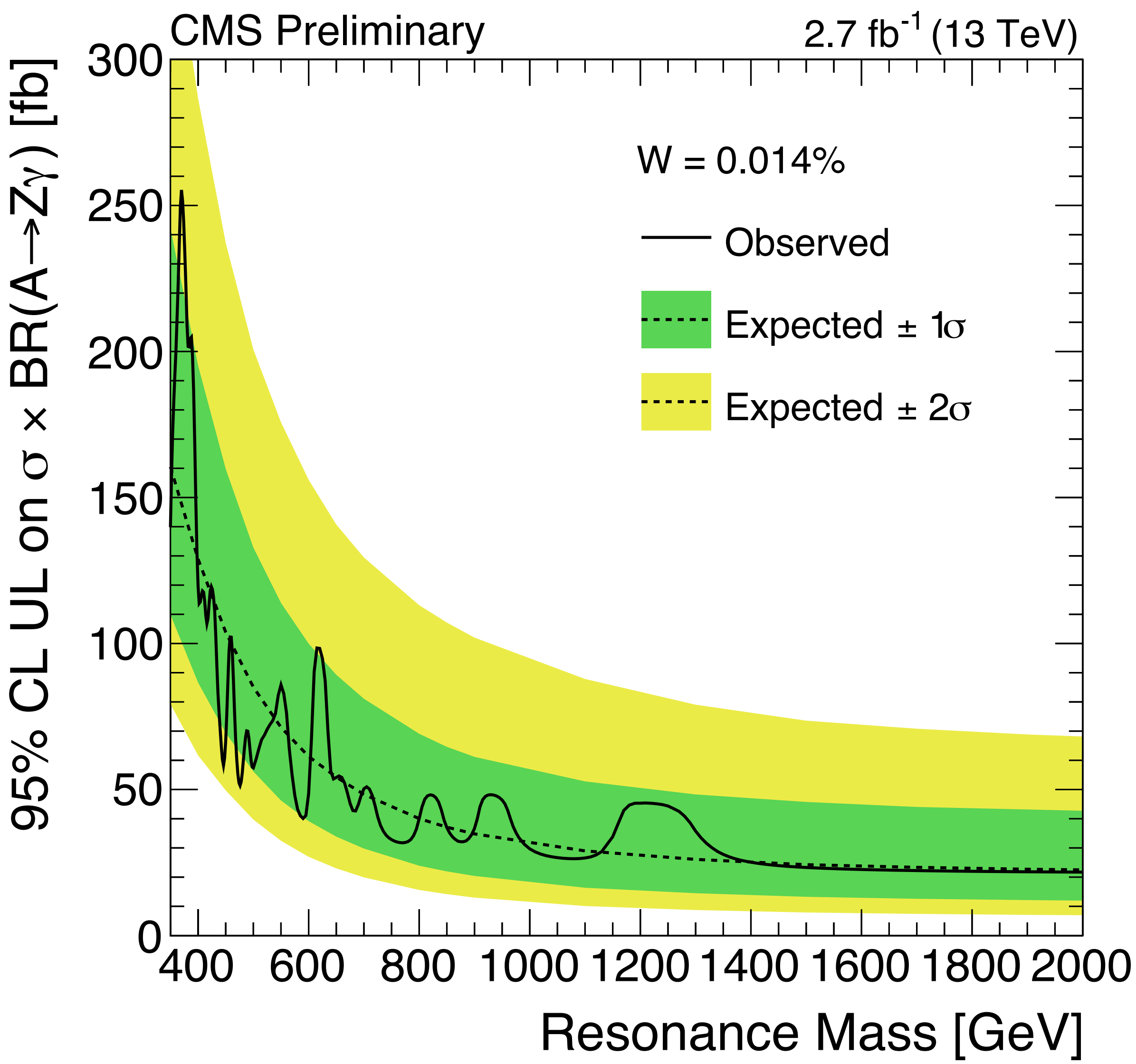
boson-tagging efficiencies

Tagger	BR(W/Z/H ➔ xx)	efficiency	mistag rate (q-/g-jets)
W/Z➔qq	70 %	35 %	1.2 %
H➔bb	57 %	35 %	0.5 %
H➔WW➔qqqq	10 %	35 %	1.5 %
H➔ττ	6 %	35 %	0.03 %

Zɣ search limits

- recently published Z➔ll + ɣ search

0.014% width



5.4% width