First experimental search for boosted Higgs boson production using the $H \rightarrow b\bar{b}$ decay with CMS

Caterina Vernieri (FNAL) on behalf of the CMS experiment

This talk

- Higgs discovery and $H \to b\bar{b}$ status at LHC
  - challenges of $H \to b\bar{b}$ at LHC
- b-tagging
  - dedicated strategy for boosted $H \to b\bar{b}$
- $H \to b\bar{b}$ tagging
- Inclusive search for boosted $H \to b\bar{b}$

Chin. Phys. C, 40, 100001 (2016)

The Higgs boson discovery

- A great advance in our understanding of fundamental particles and their interactions
  - a completely **new state** of matter-energy
  - properties as a **potential window to Beyond Standard Model (SM)**

H^0

$J = 0$

Mass $m = 125.09 \pm 0.24$ GeV
Full width $\Gamma < 1.7$ GeV, CL = 95%

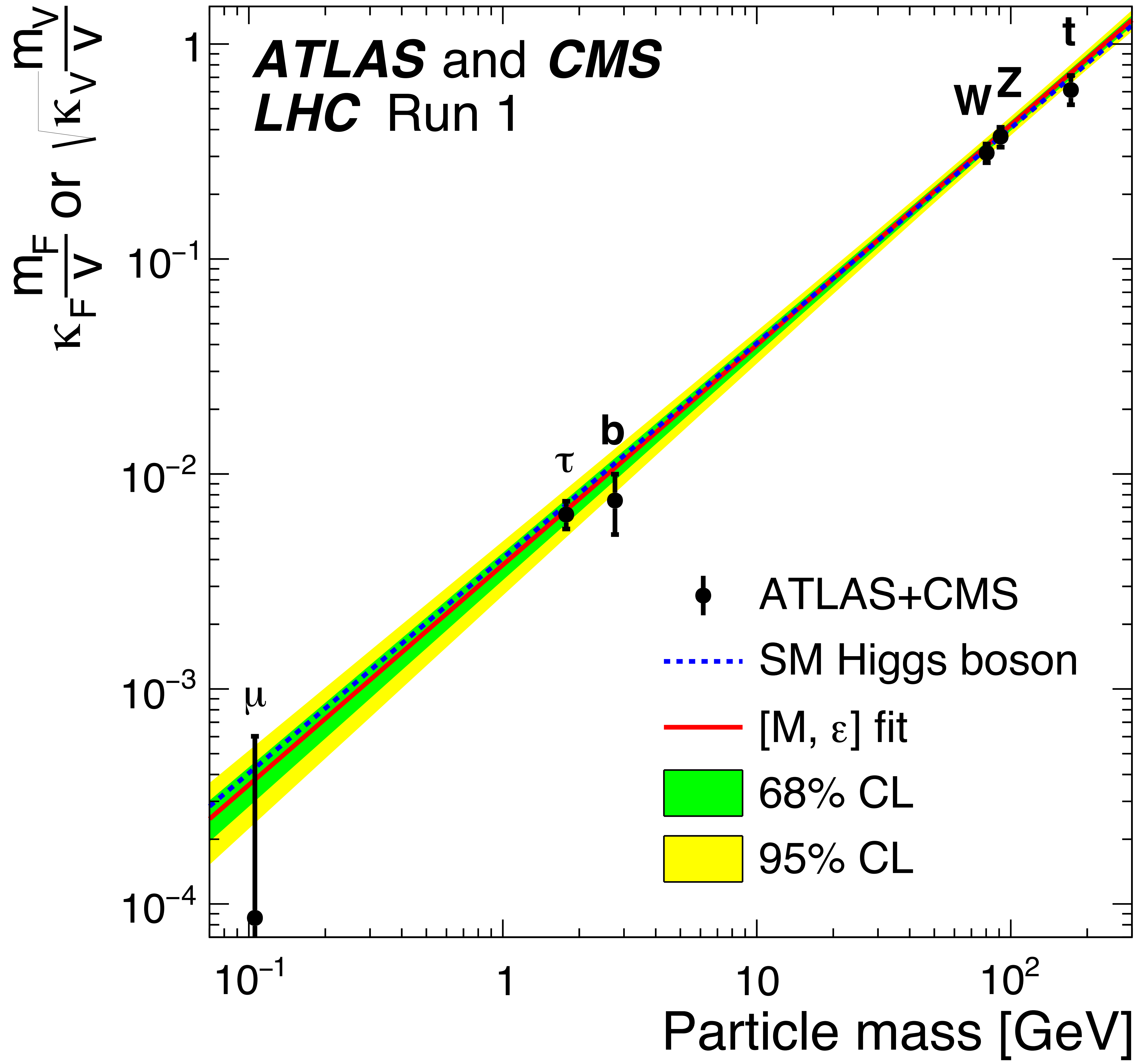
H^0 Signal Strengths in Different Channels

See Listings for the latest unpublished results.

Combined Final States $= 1.10 \pm 0.11$
$WW^* = 1.08^{+0.18}_{-0.16}$
$ZZ^* = 1.29^{+0.26}_{-0.23}$
$\gamma\gamma = 1.16 \pm 0.18$
$b\bar{b} = 0.82 \pm 0.30$ (S = 1.1)
$\mu^+\mu^- < 7.0$, CL = 95%
$\tau^+\tau^- = 1.12 \pm 0.23$
$Z\gamma < 9.5$, CL = 95%
$t\bar{t}H^0$ Production $= 2.3^{+0.7}_{-0.6}$

H to $b\bar{b}$

- Observed production and decay rates are **consistent with a SM Higgs boson** within uncertainties
- Unique final state to study the **coupling with down-type quark**
- **Largest BR** for SM H (~58%), dominates total width
    - not yet observed

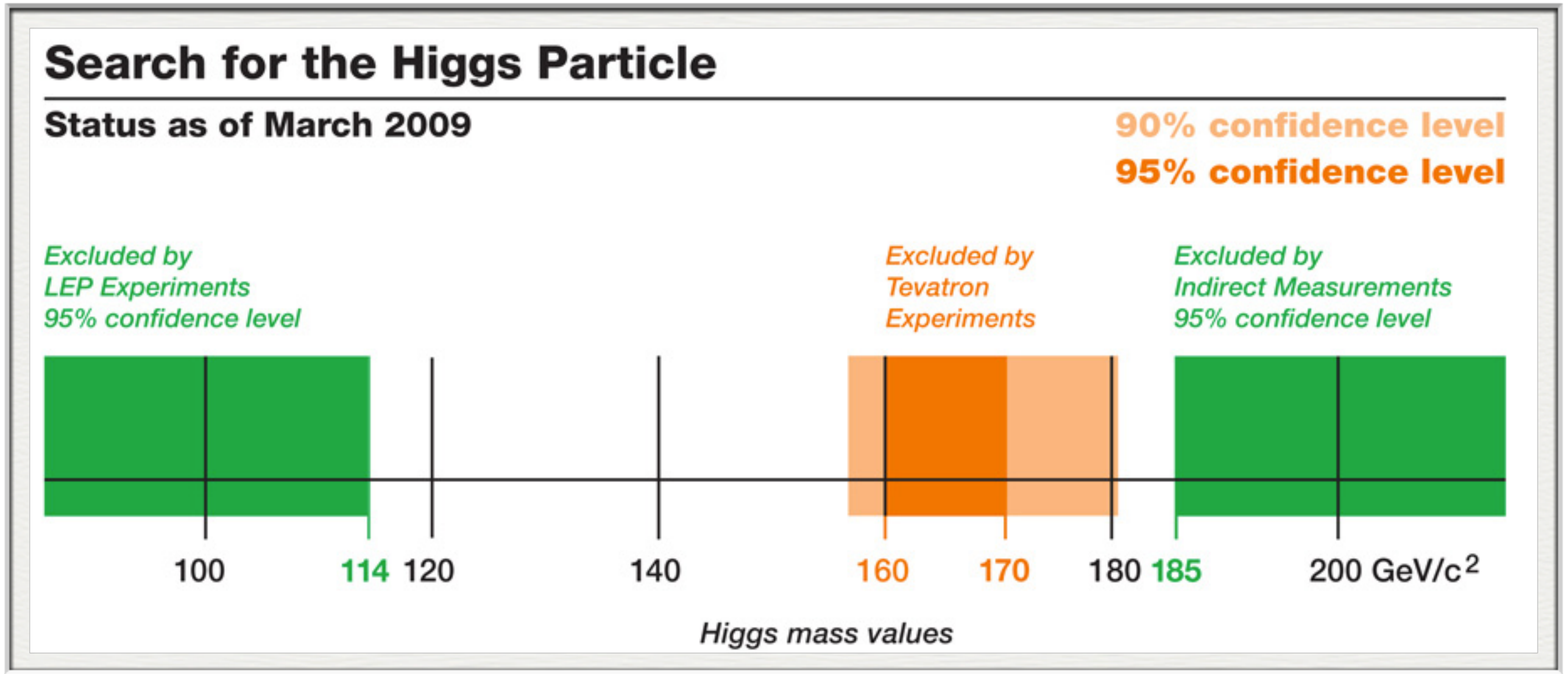


Standard Model Higgs Hunting: Basics

LEP+Tevatron legacy:

low-mass range [114,158] GeV

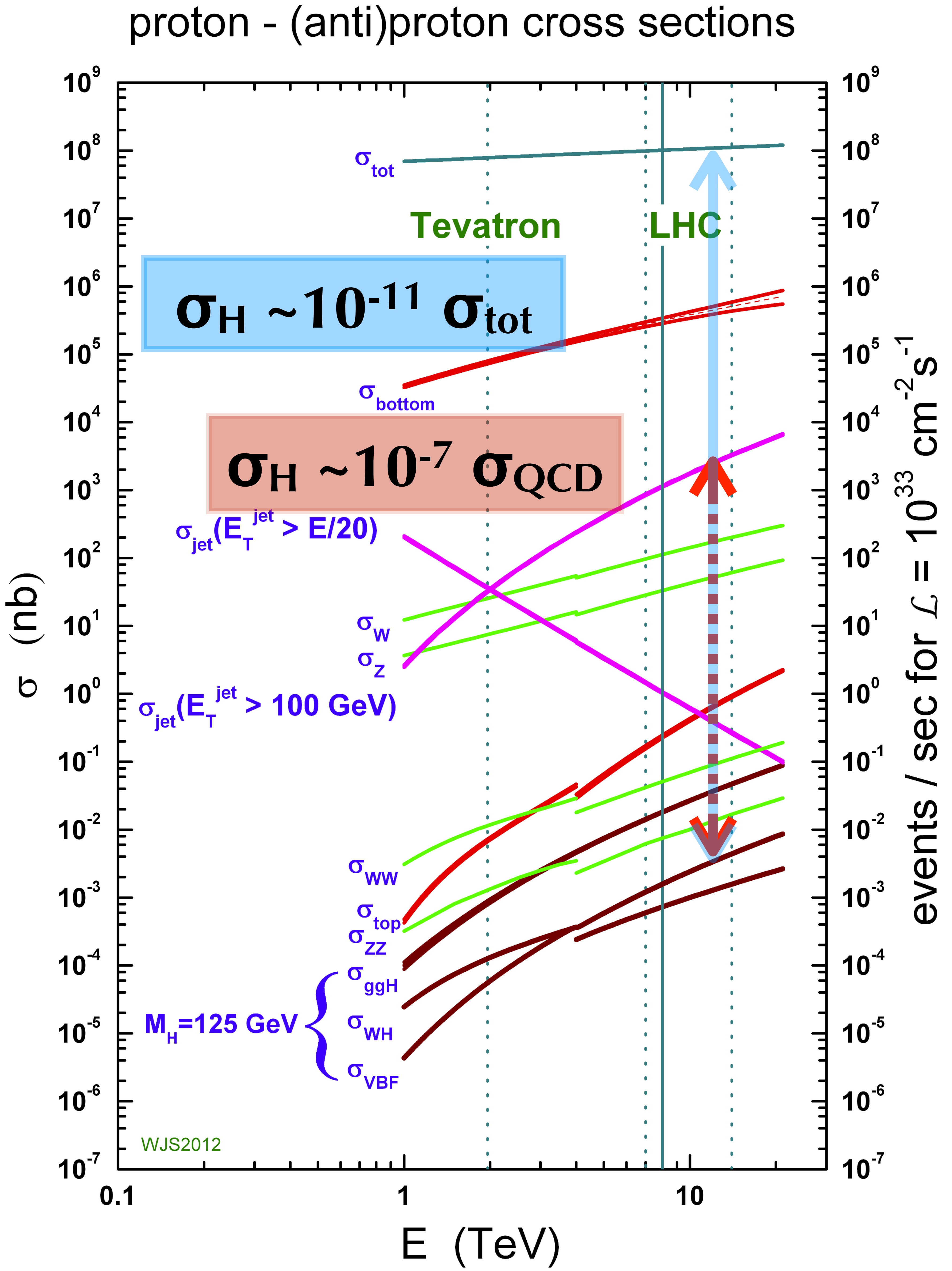
2009



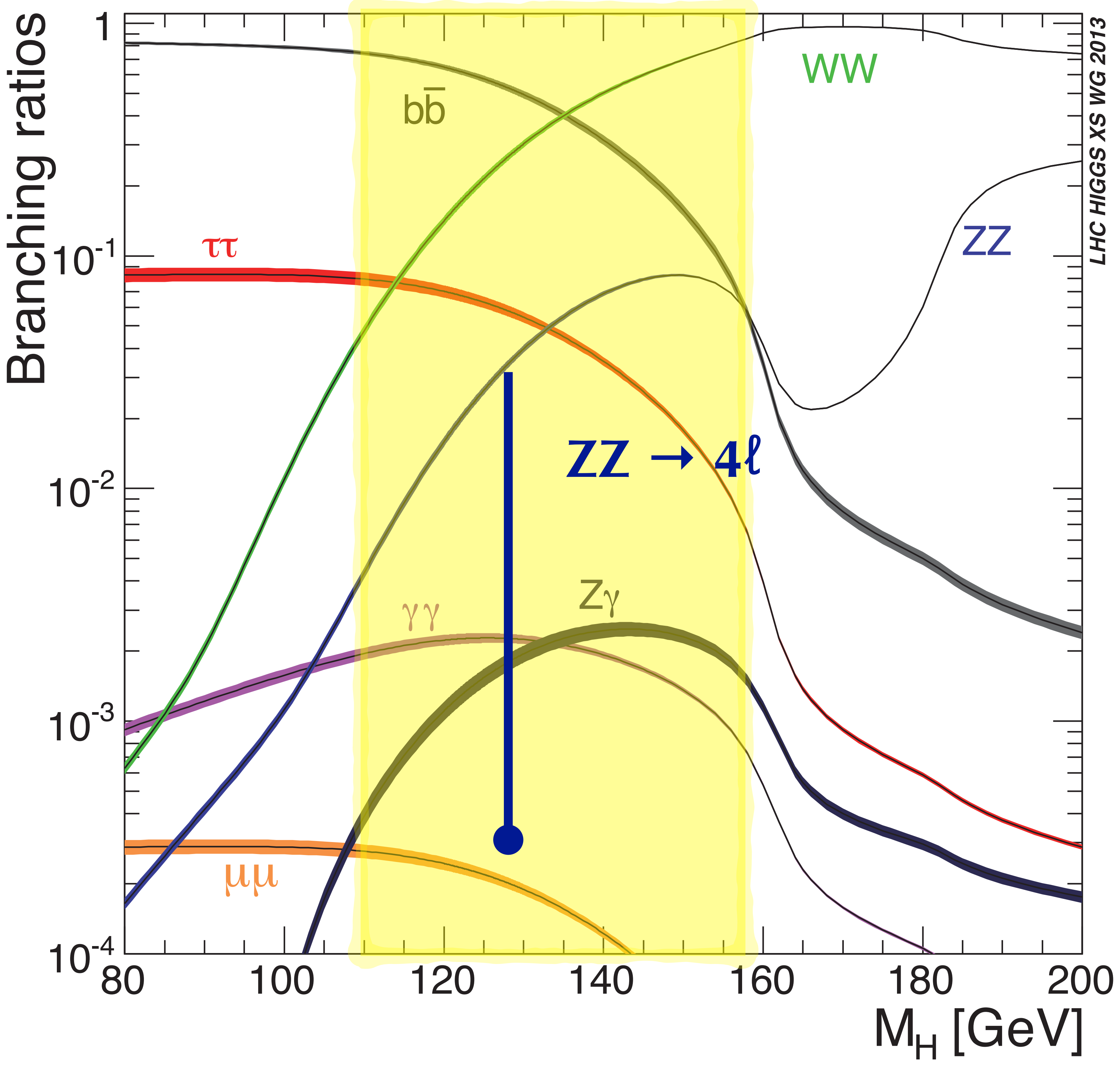
The natural width is less than 100 MeV

observed peak dominated by instrumental mass resolution

Standard Model Higgs Hunting: Strategy



H needs
- high energy
- optimal invariant mass resolution
- luminosity

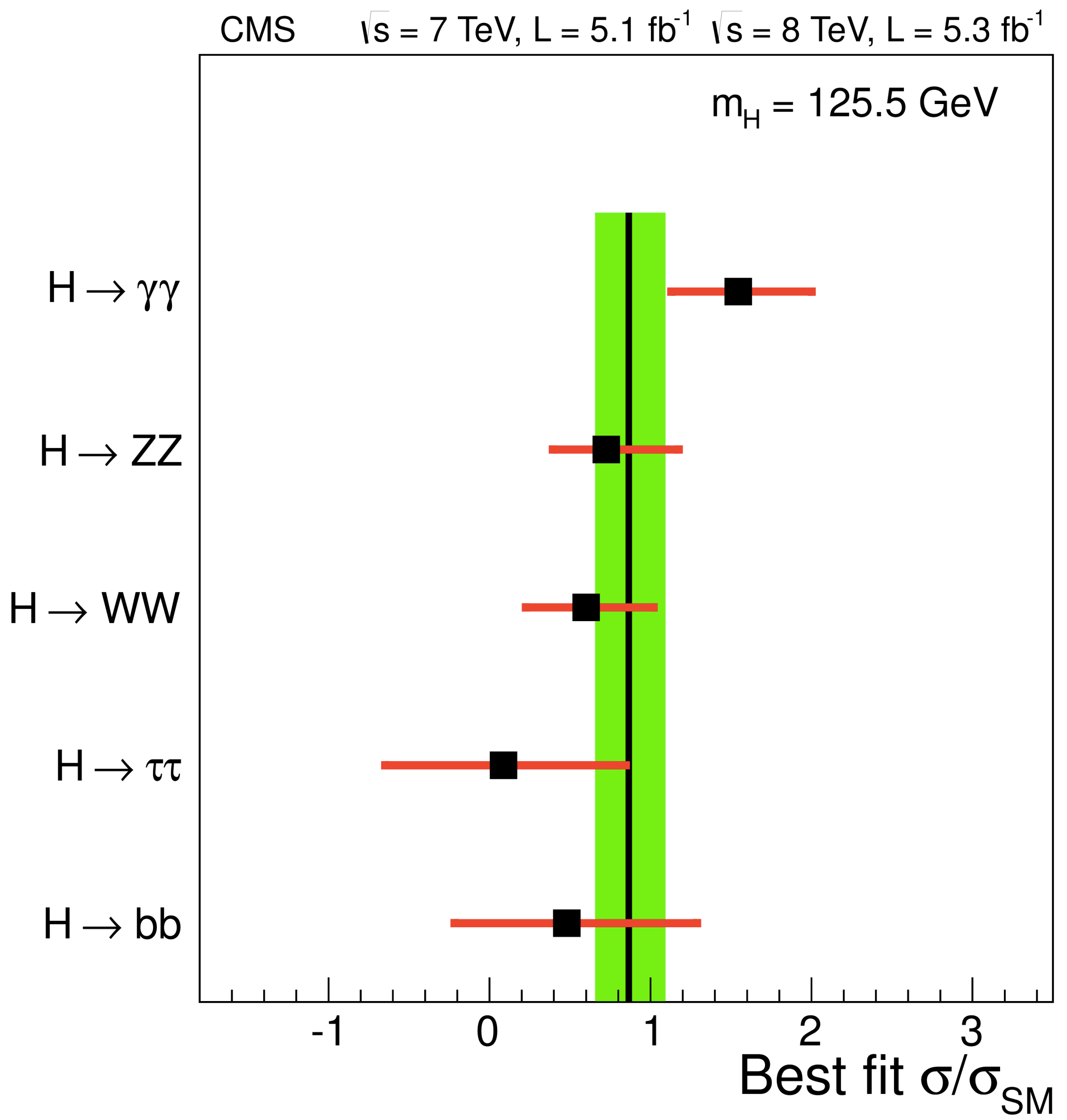


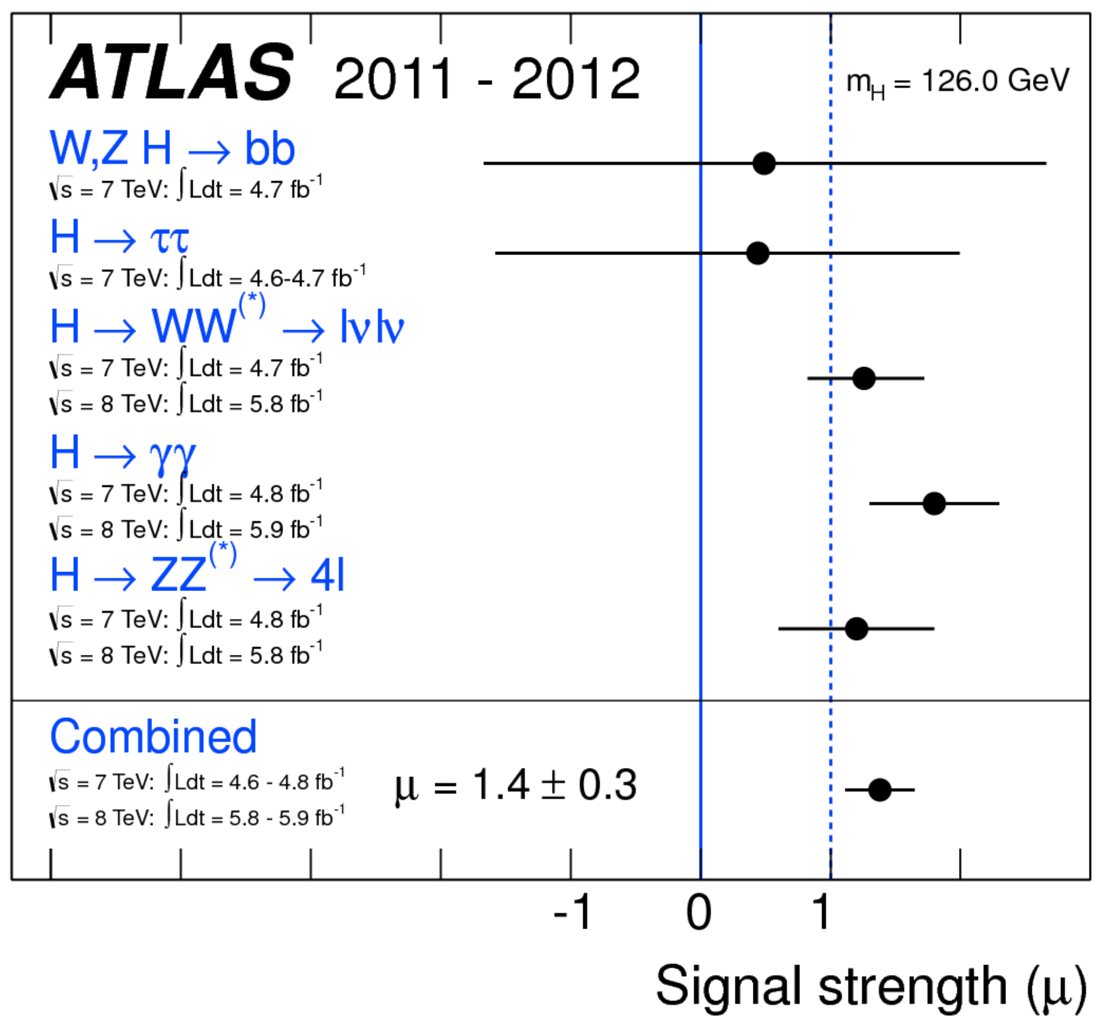
lowest BR decay modes ↔ Excellent mass resolution

Phys. Lett. B 716 (2012) 1
Phys. Lett. B 716 (2012) 30

July 4th 2012, H discovery

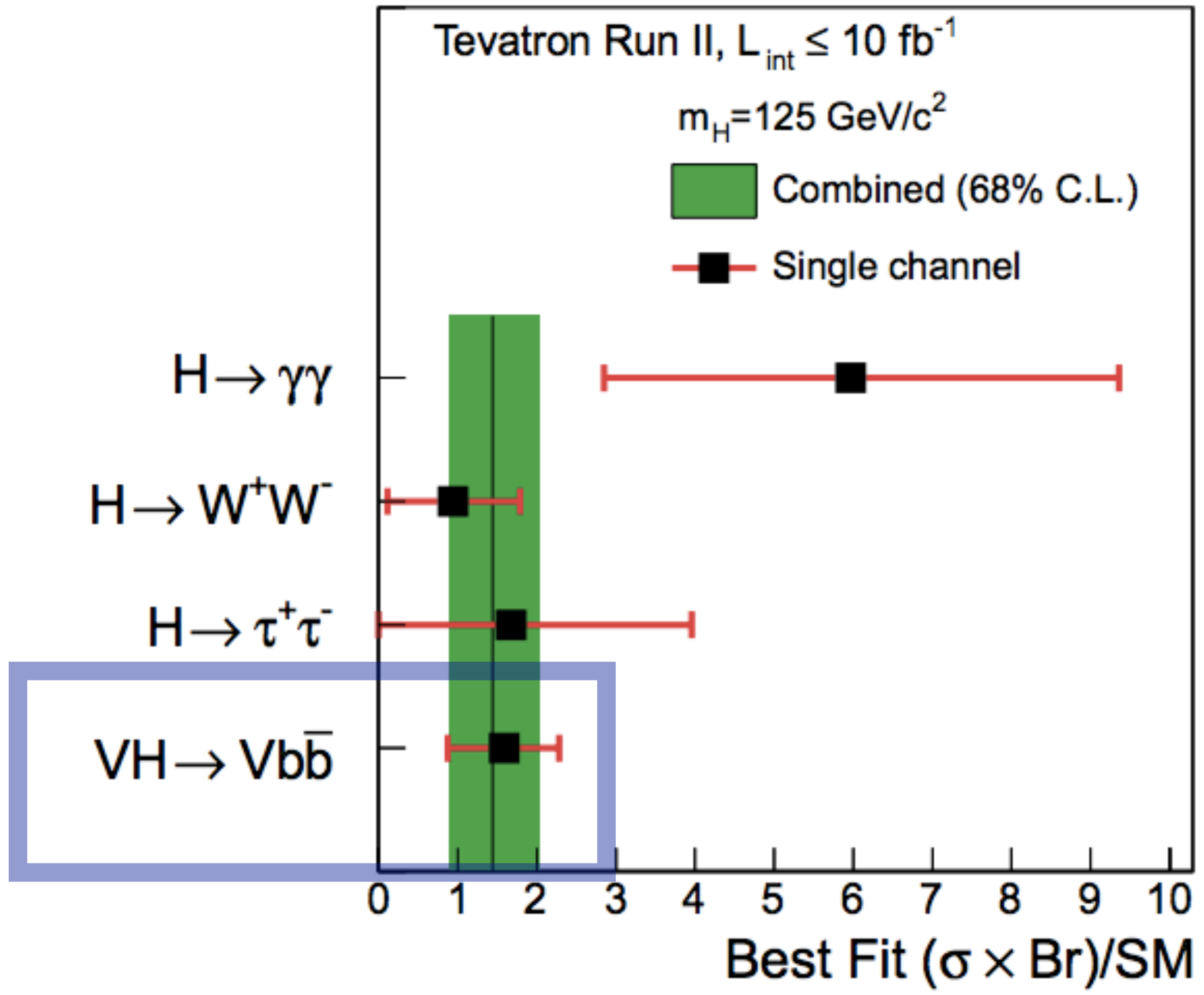
- CMS and ATLAS reported independently the **first observation** of the Higgs boson
  - **5.0 σ combining γγ and ZZ alone**
    - **best mass resolution**
    - the huge amount of LHC data allows to exploit the lowest BR decay modes





- CDF and D0 combined results reported a broad excess in the mass range $115<m_H<140$ GeV
  - 3.0 σ at m_H=125 GeV
  - **mainly from the H → $b\bar{b}$**

CDF+D0, 2.8 σ

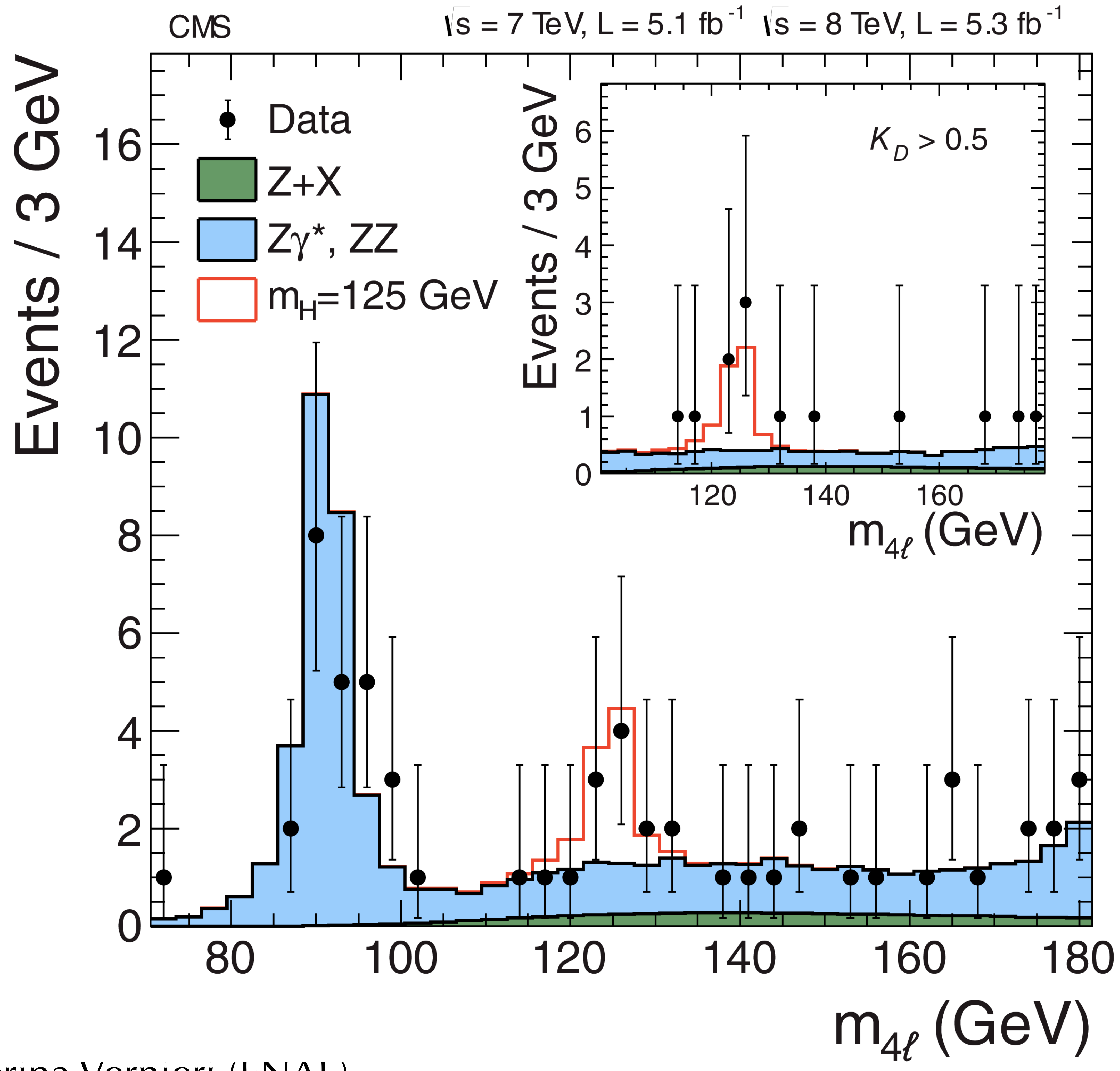


arXiv:1303.6346

Phys. Lett. B 716 (2012) 30
Phys. Rev. D 89 (2014)

Challenges of the H($b\bar{b}$) mode at the LHC

Comparison with one of the discovery channels

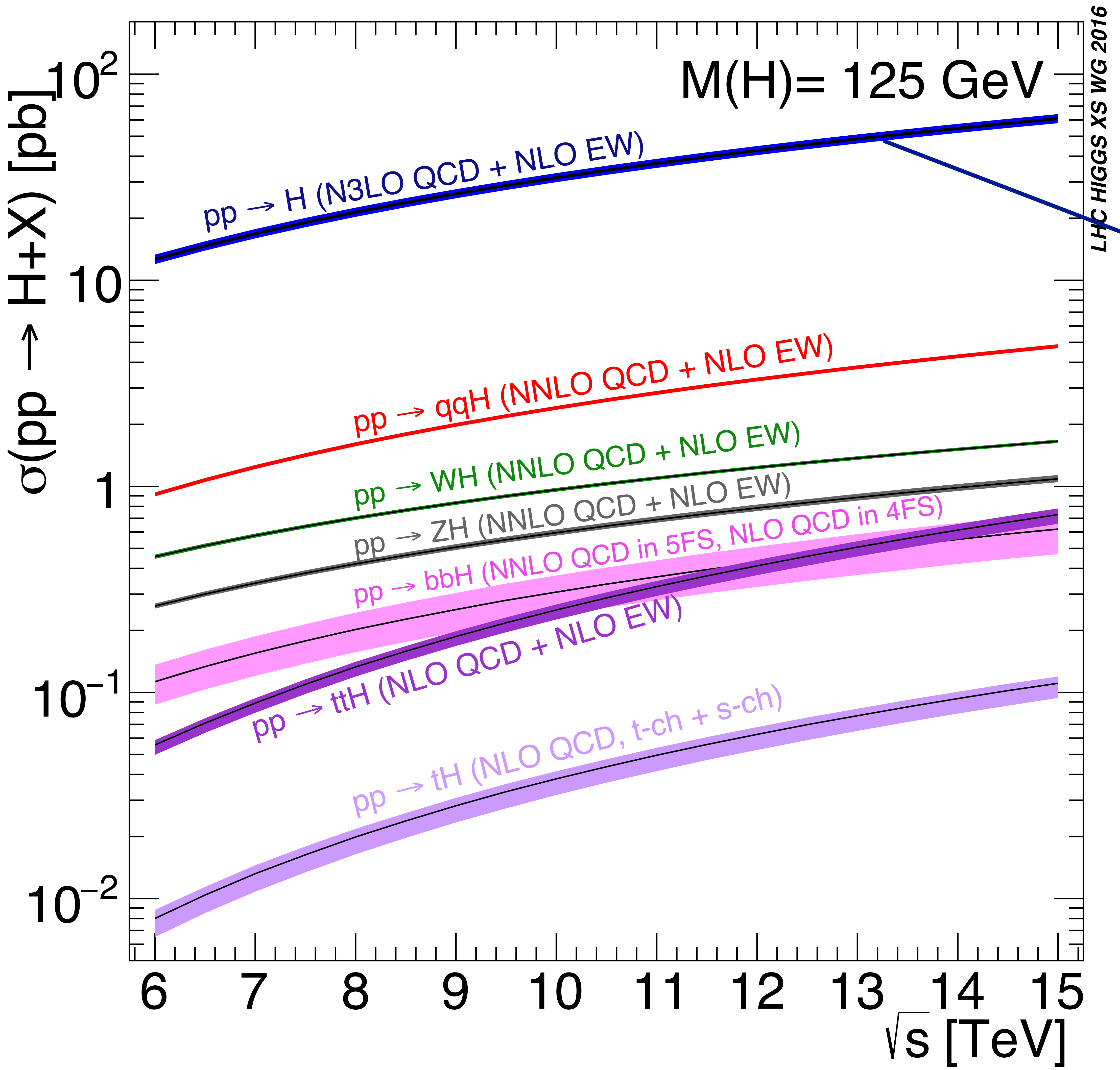


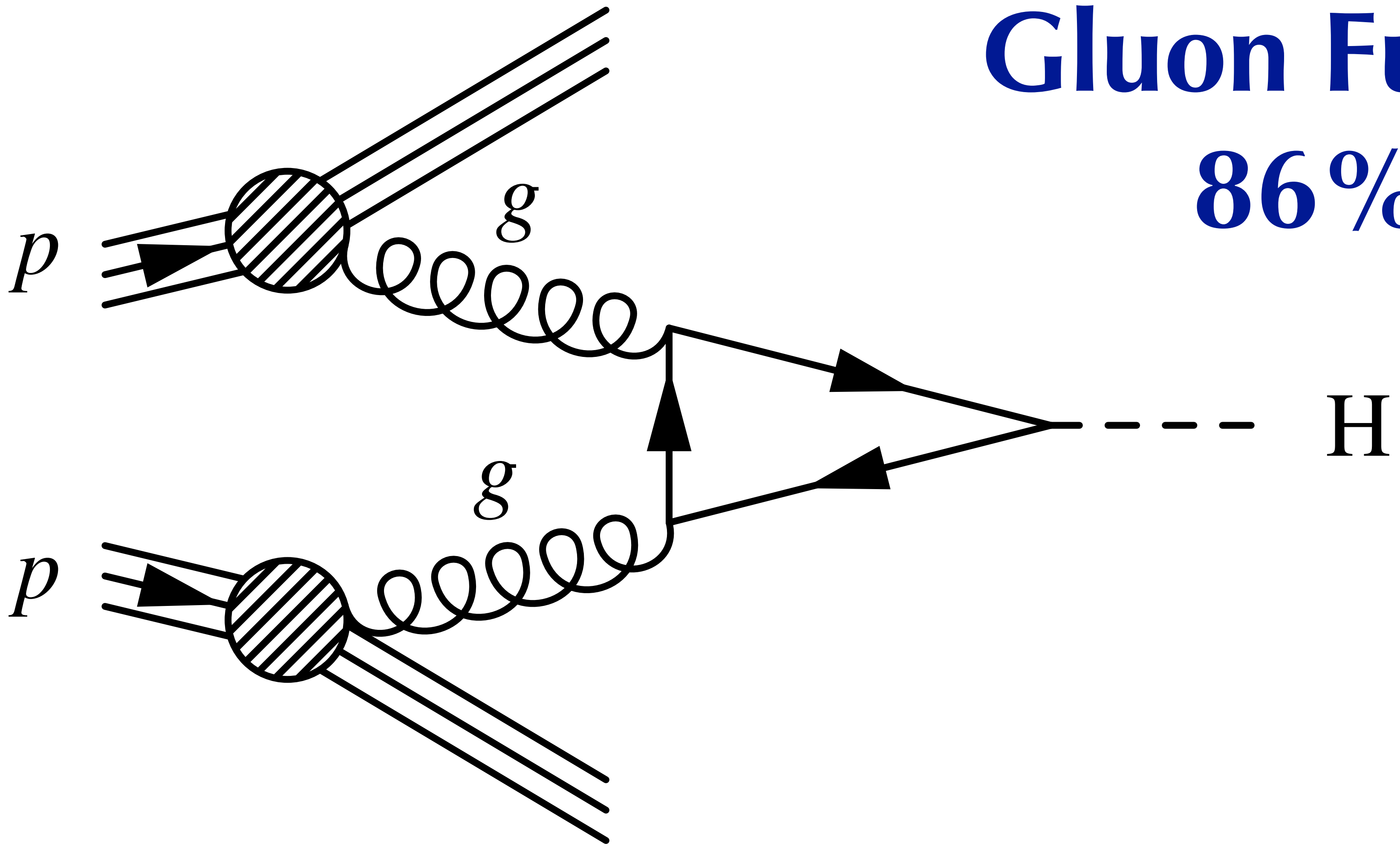
	$H \to 4\ell$	$H \to b\bar{b}$
BR	0.03%	58%
mass resolution	1%	10%
signal efficiency	30%	1.3%
S/B	2	0.05

H($b\bar{b}$) searches need:

- b-jets identification
- improve m($b\bar{b}$) resolution
- exploit all possible information from the event to improve S/B

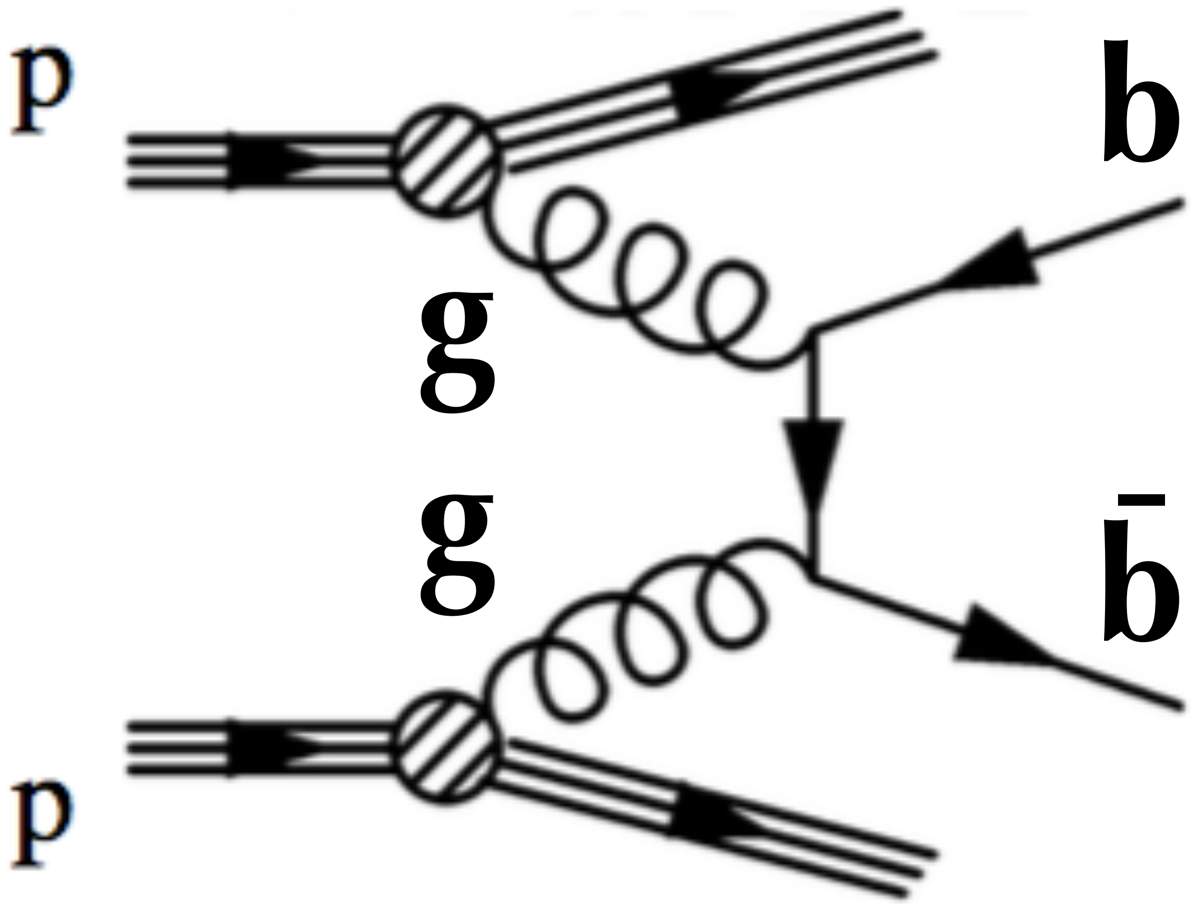
H → $b\bar{b}$ at LHC

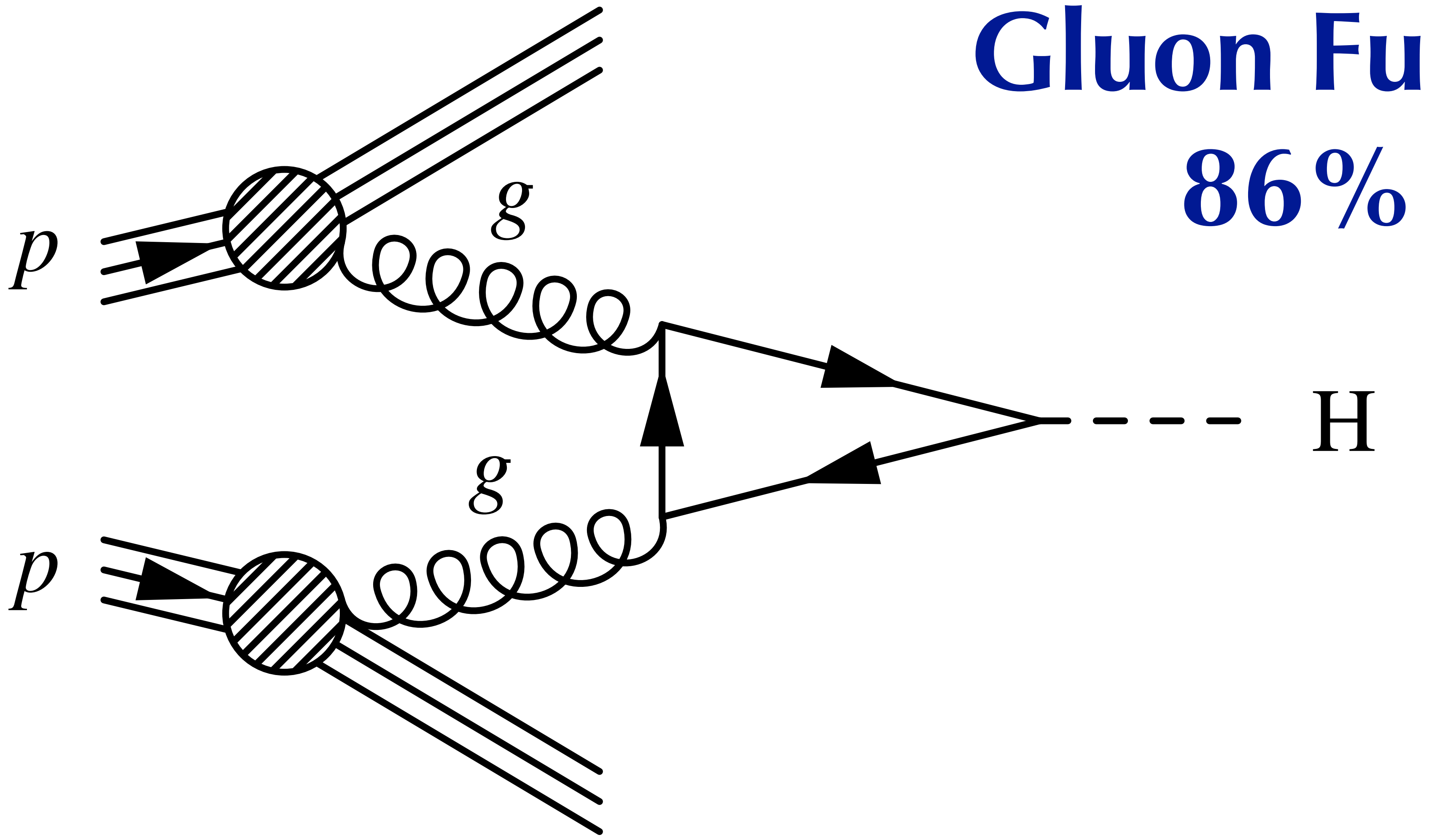




Gluon Fusion
86%

$H \rightarrow b\bar{b}$ at LHC





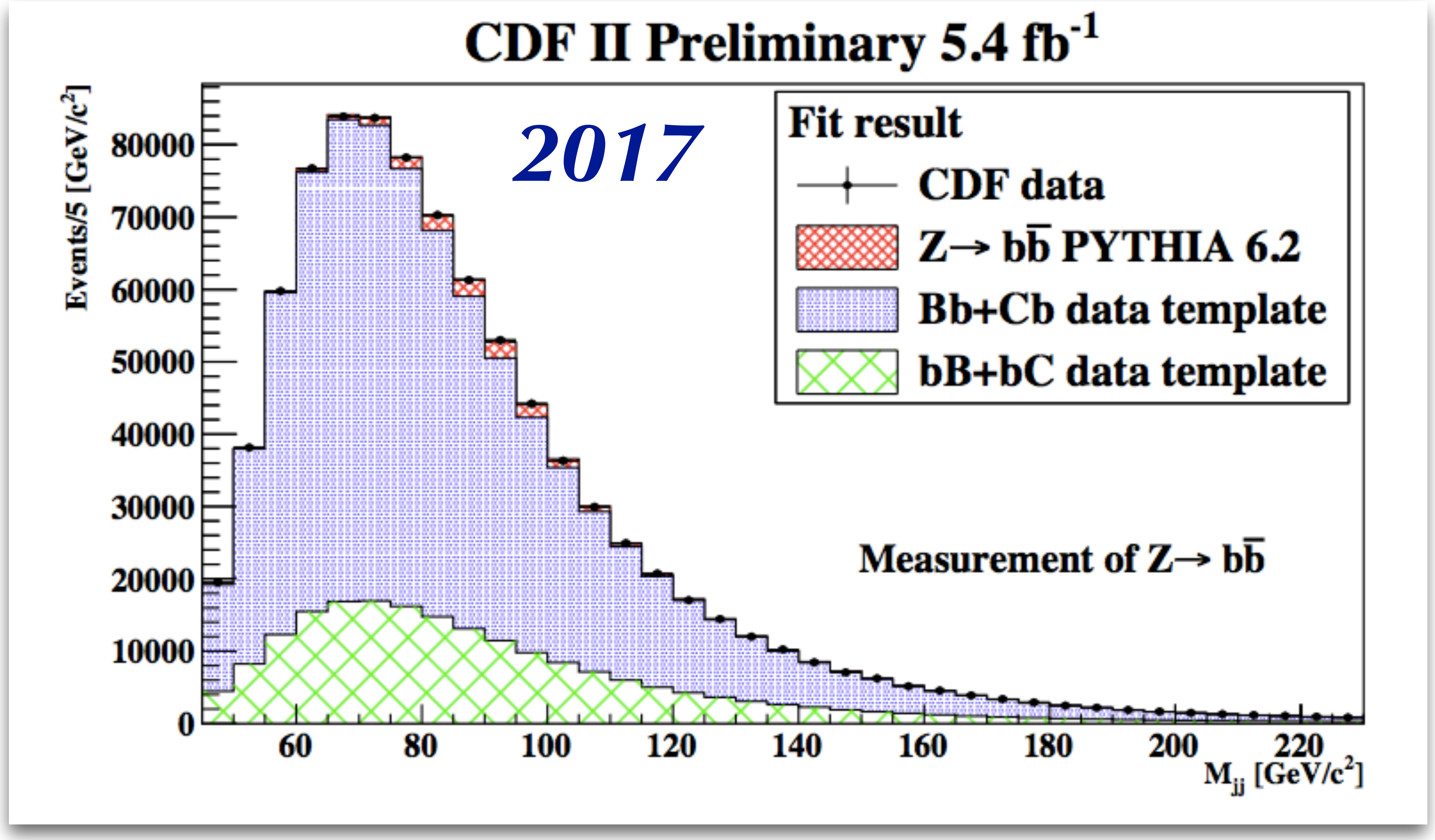
Gluon Fusion 86%

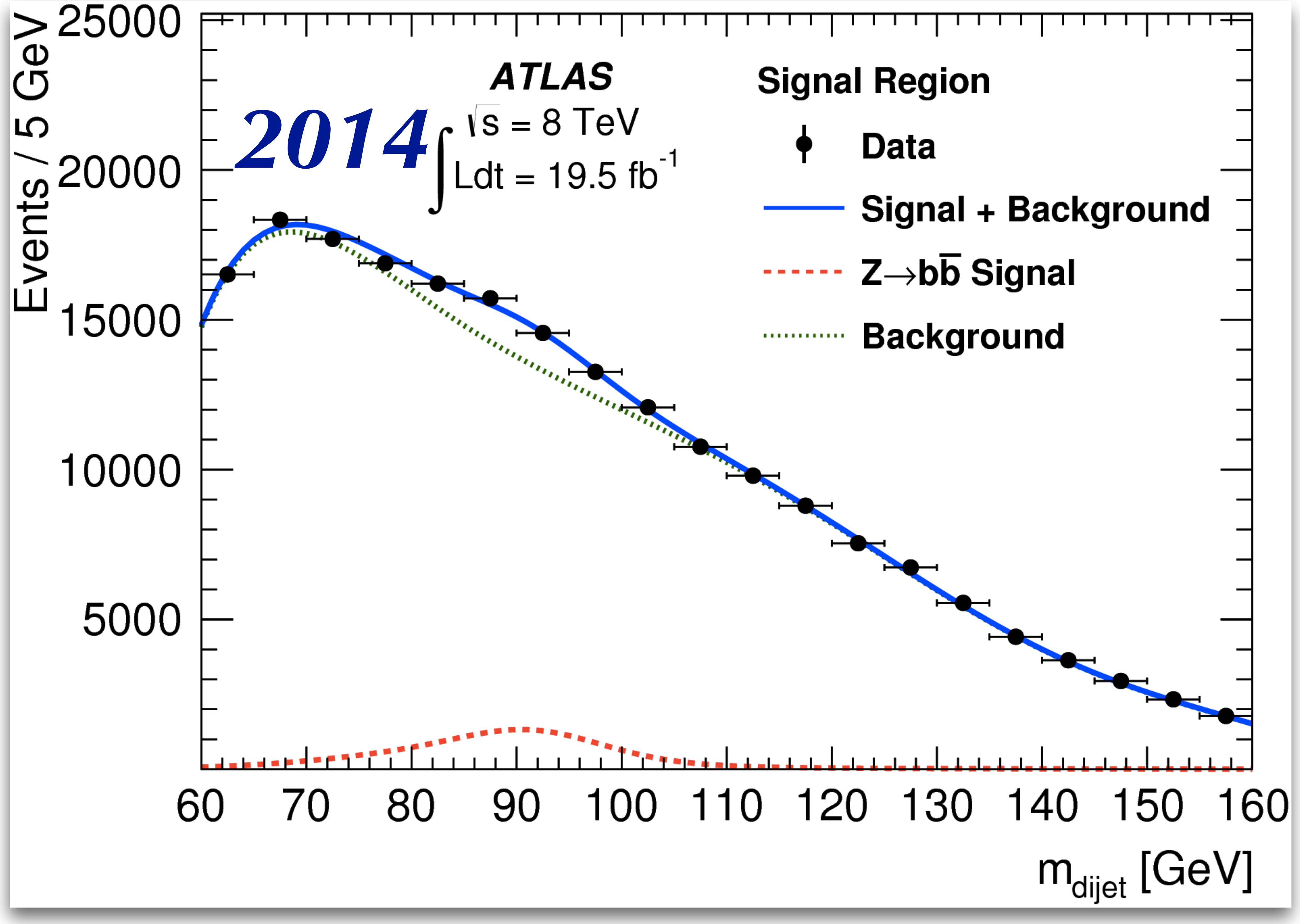
- Overwhelming background from QCD production of b quarks
  - 10^7 larger

Search for inclusive H to $b\bar{b}$

CDF-11228
Phys. Lett. B 738 (2014)

Two-jets final state, overwhelming multijets background

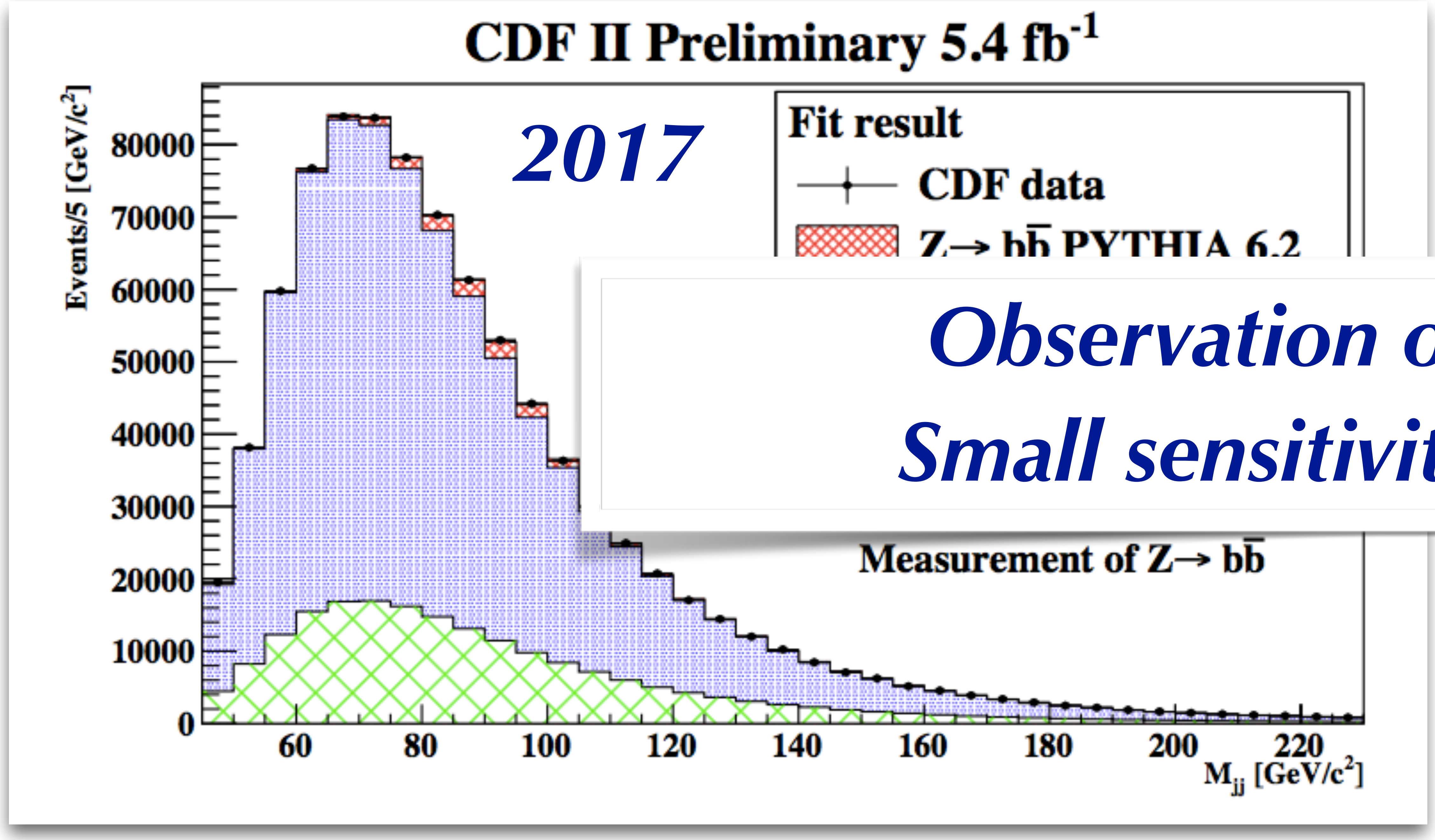


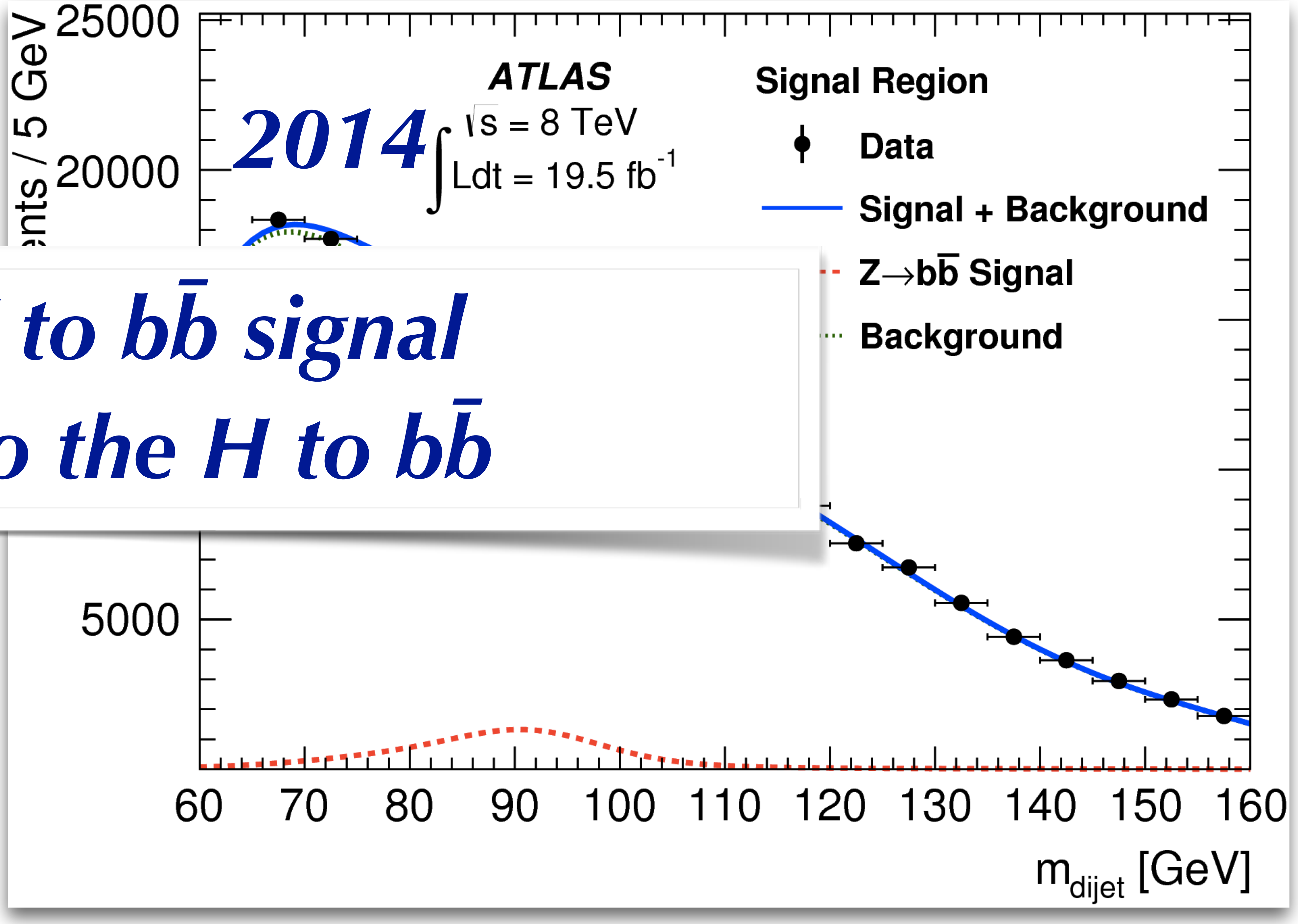


Search for inclusive H to $b\bar{b}$

CDF-11228
Phys. Lett. B 738 (2014)

Two-jets final state, overwhelming multijets background

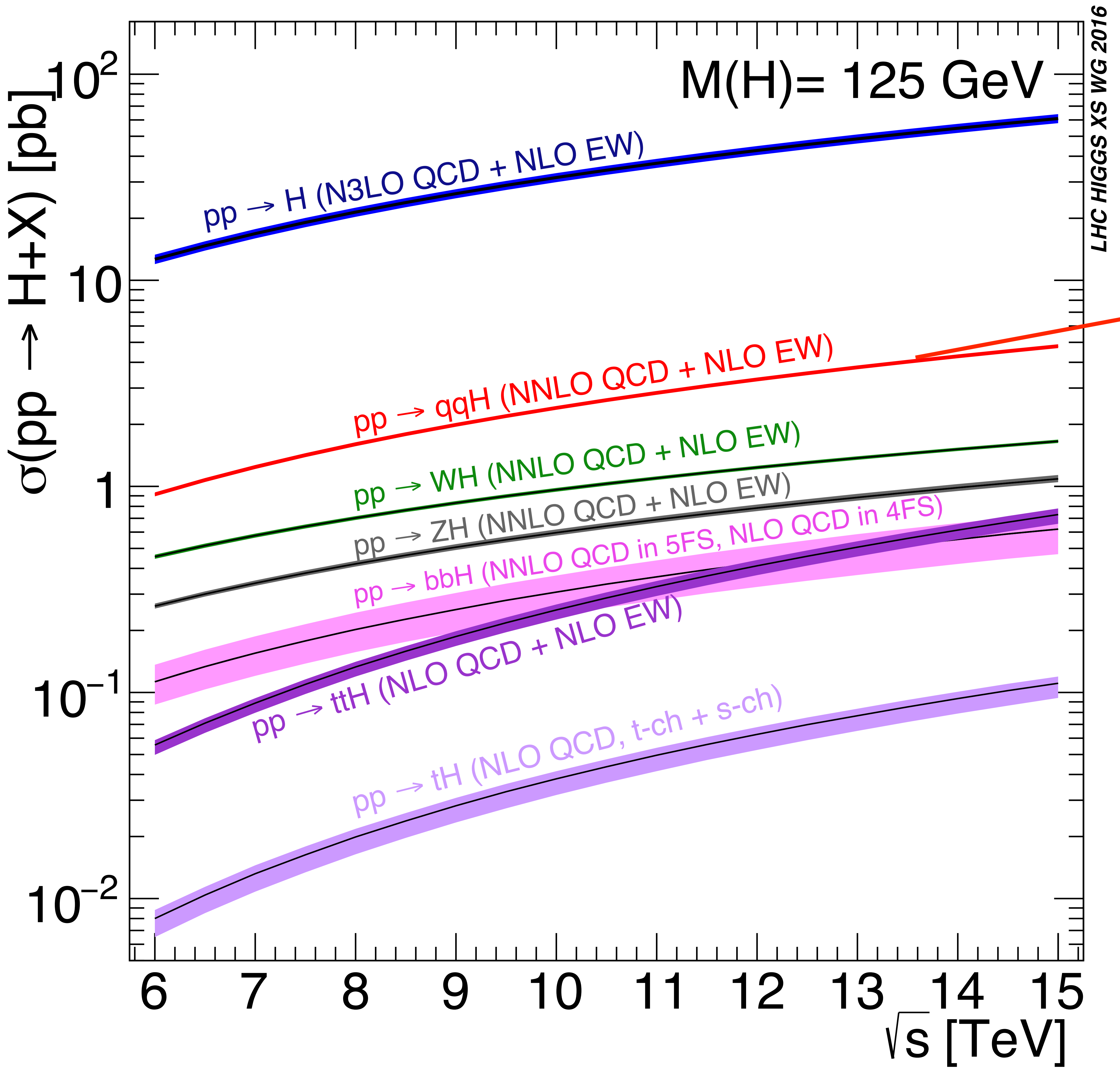


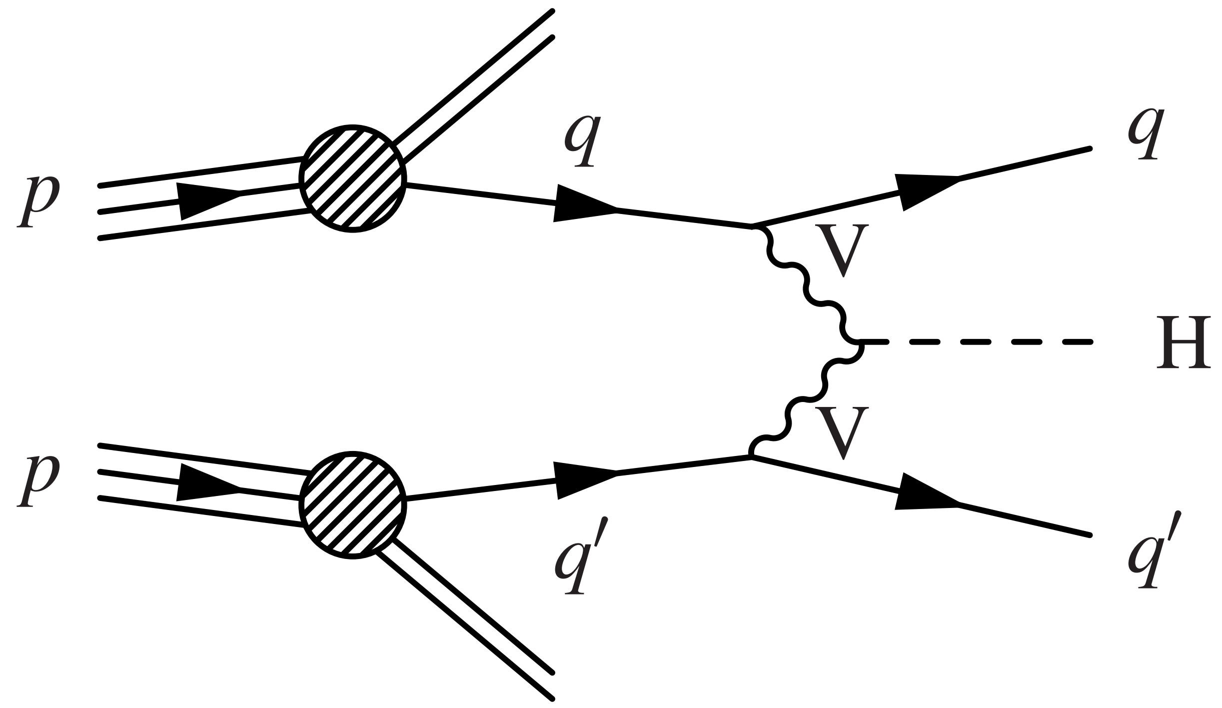


Observation of Z to $b\bar{b}$ signal
Small sensitivity to the H to $b\bar{b}$

Search for gg → H → $b\bar{b}$ historically **deemed impossible**

$H \to b\bar{b}$ at LHC



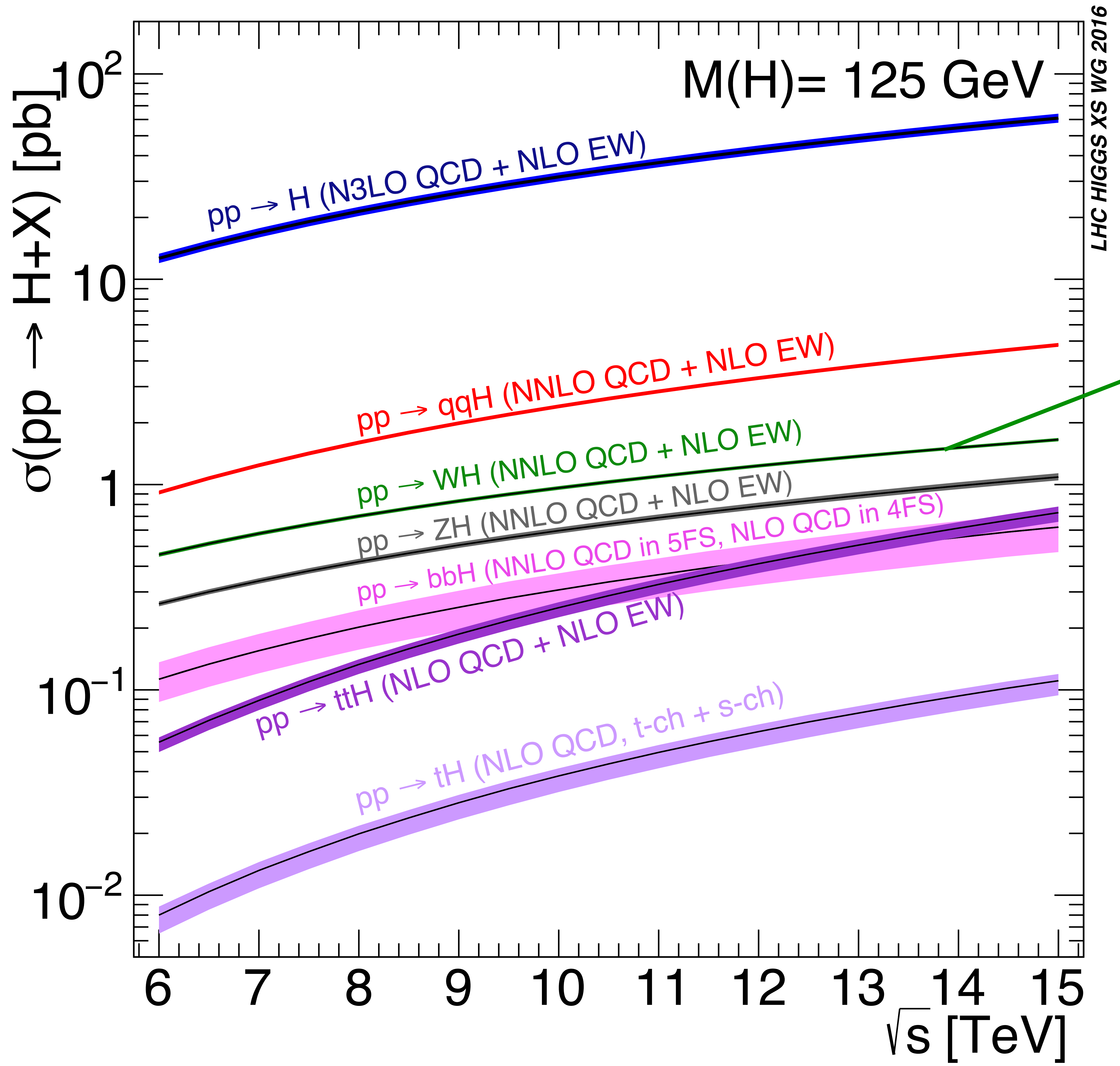


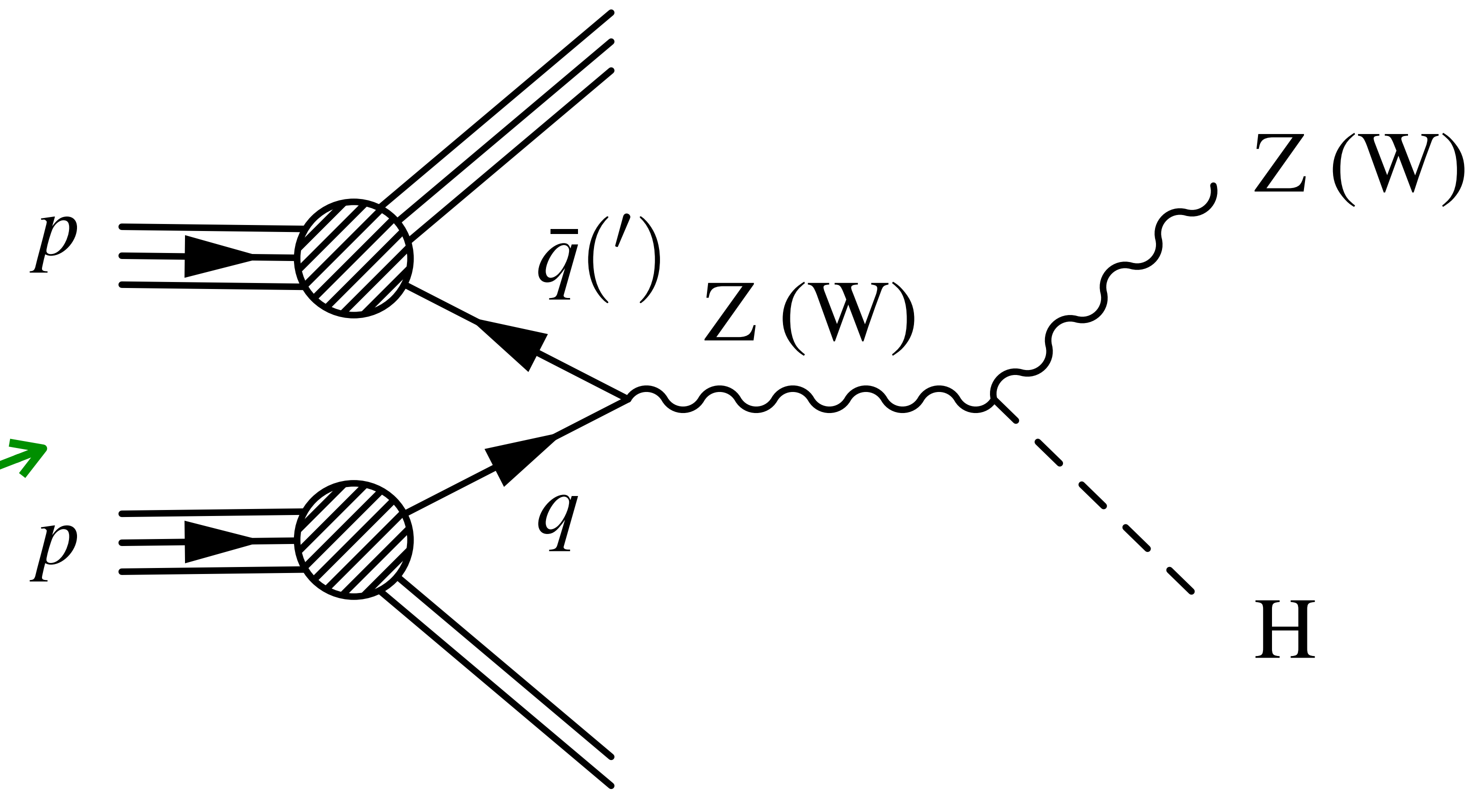
VBF
7%

large background but a very distinctive topology

Caterina Vernieri (FNAL)

$H \to b\bar{b}$ at LHC



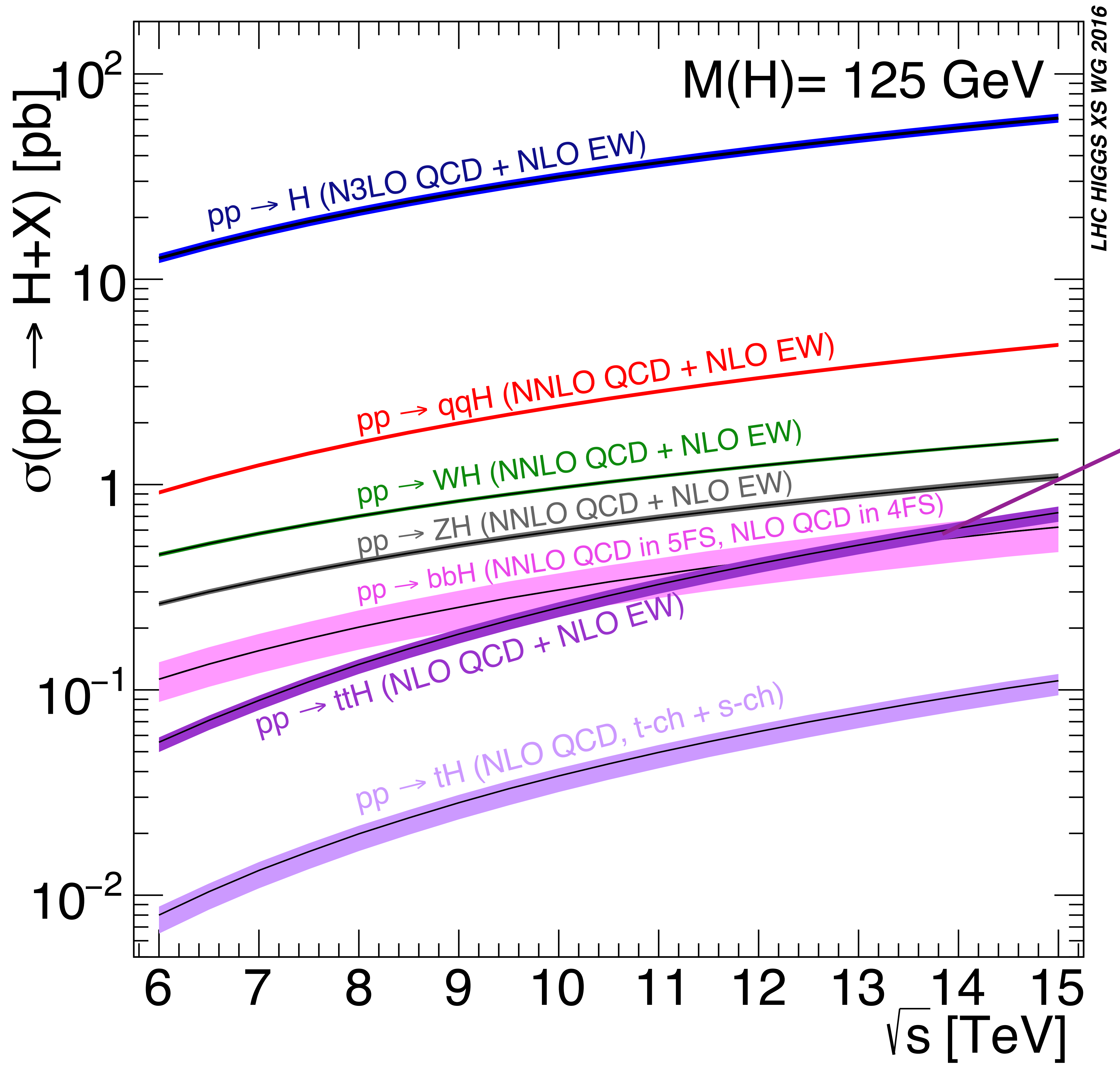


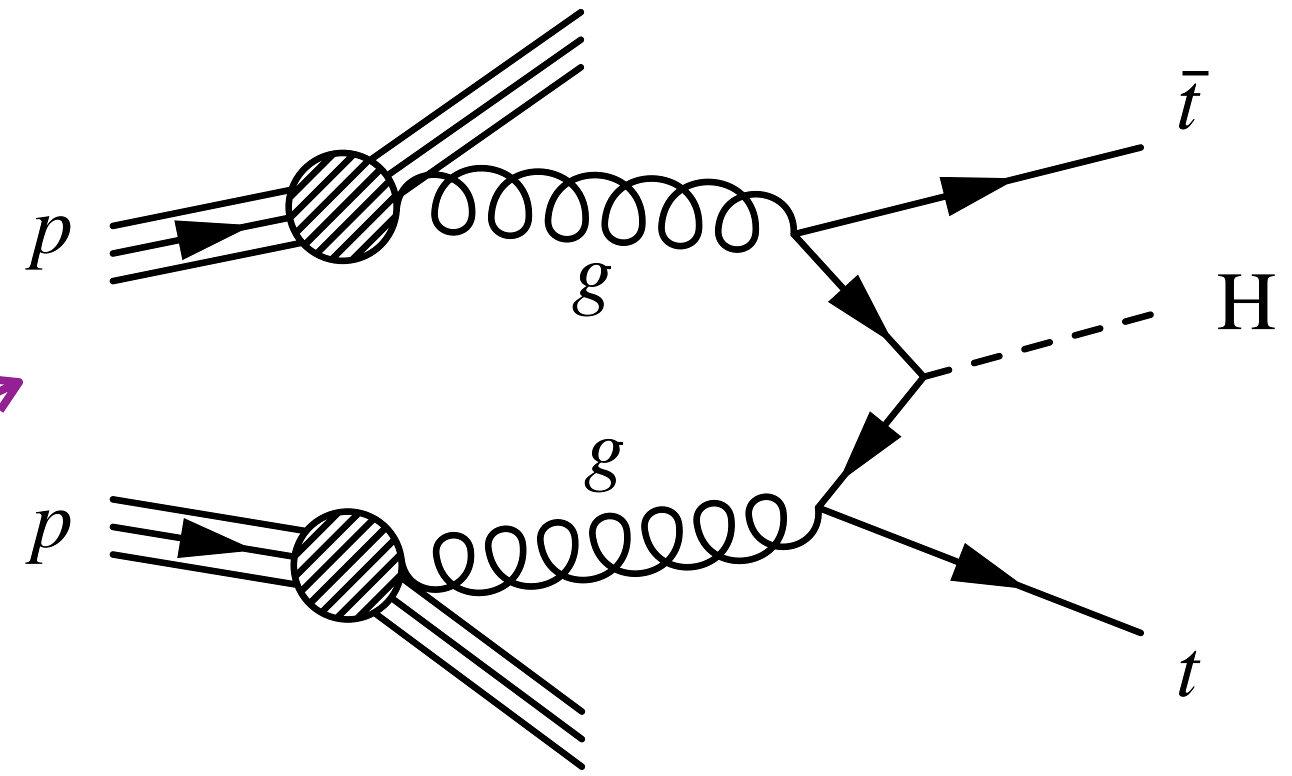
VH

4%

- W/Z decaying leptonically
  - leptons, E_T^{mis} to trigger and suppress backgrounds
- Requiring high p_T V to reduce multijets background
  - S/B at LHC is 2.5x lower than at Tevatron

$H \to b\bar{b}$ at LHC





ttH
1%

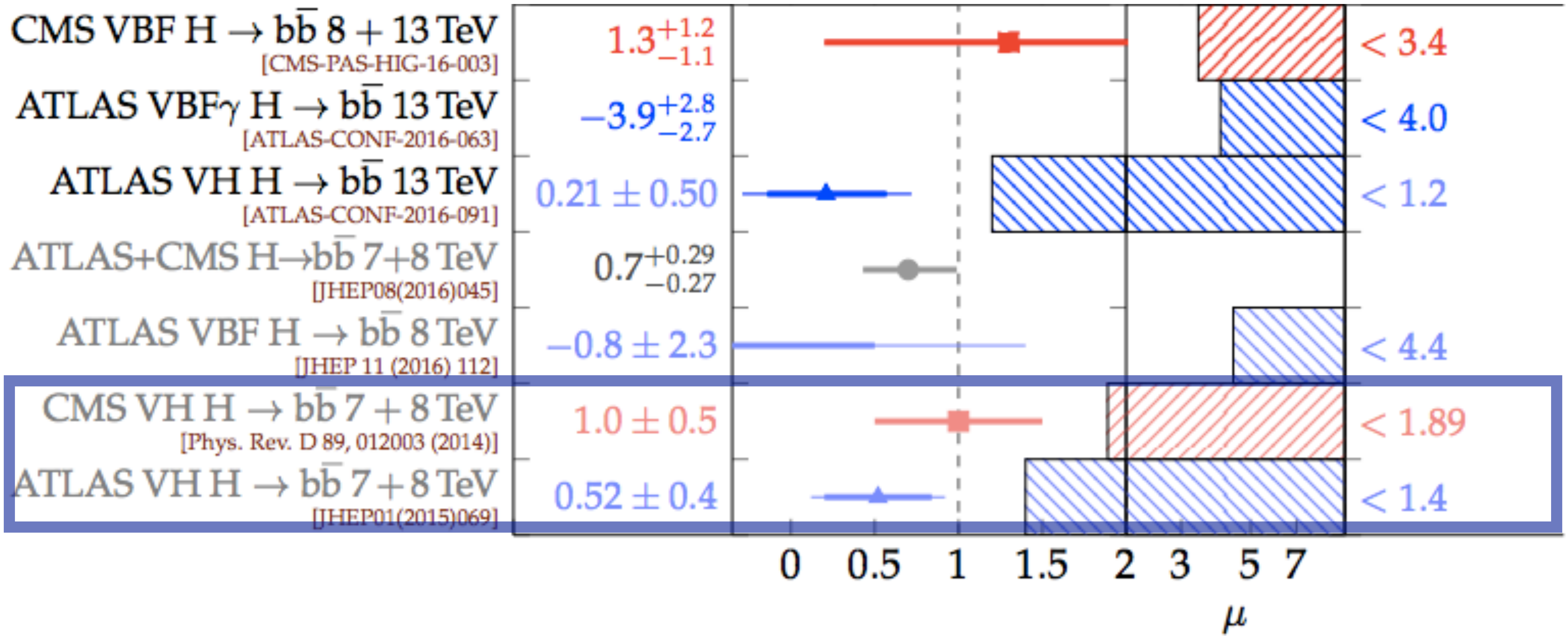
- direct probe of the top-Higgs coupling
- dominant backgrounds is $t\bar{t}$ + jets

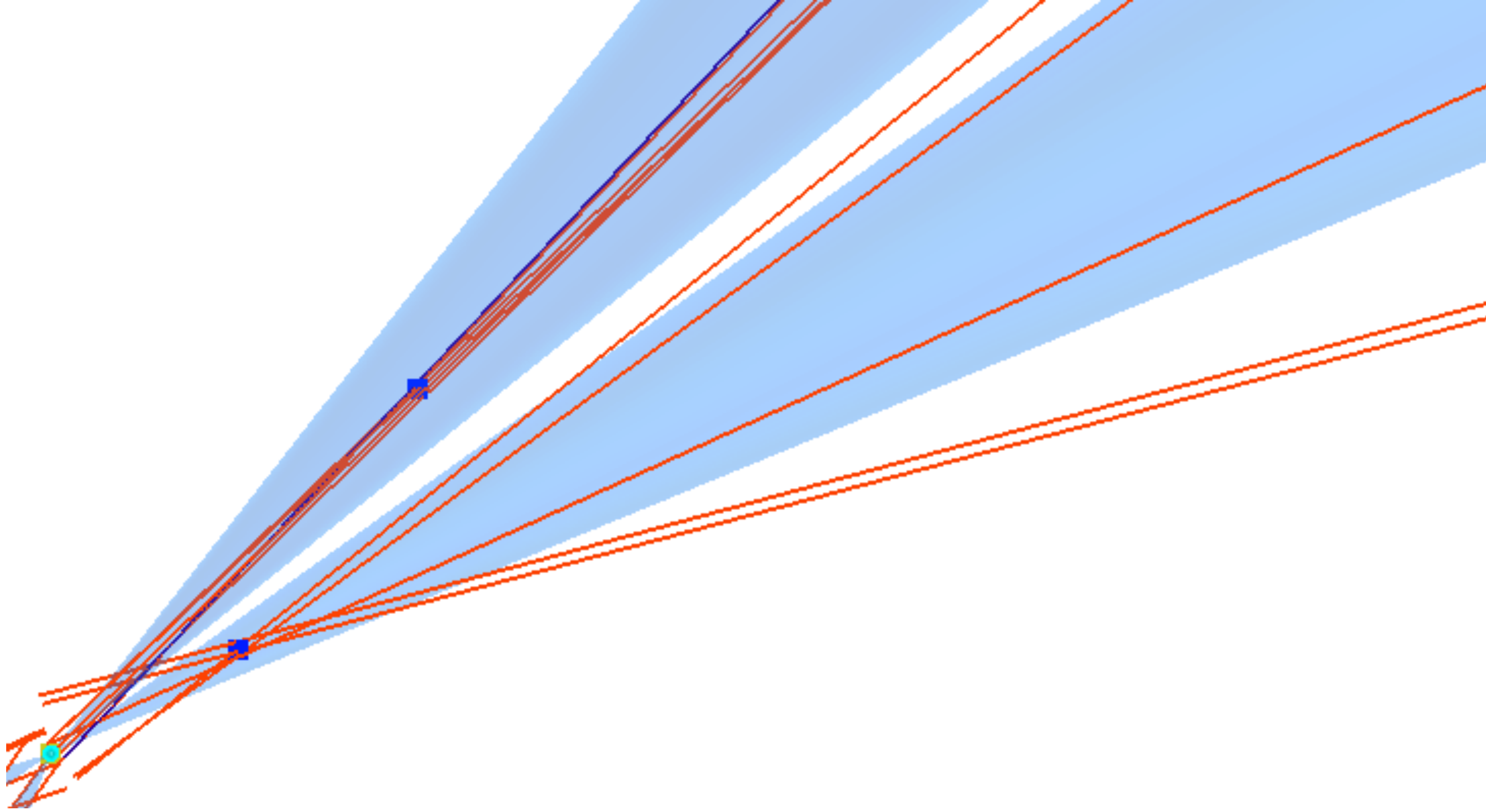
Status of H($b\bar{b}$) at LHC

The LHC combination of the Run 1 ATLAS and CMS analyses resulted in a significance of **2.6σ** (3.7σ) **observed** (expected)

VH is the **most sensitive channel for H($b\bar{b}$)**

courtesy of G. Gaycken



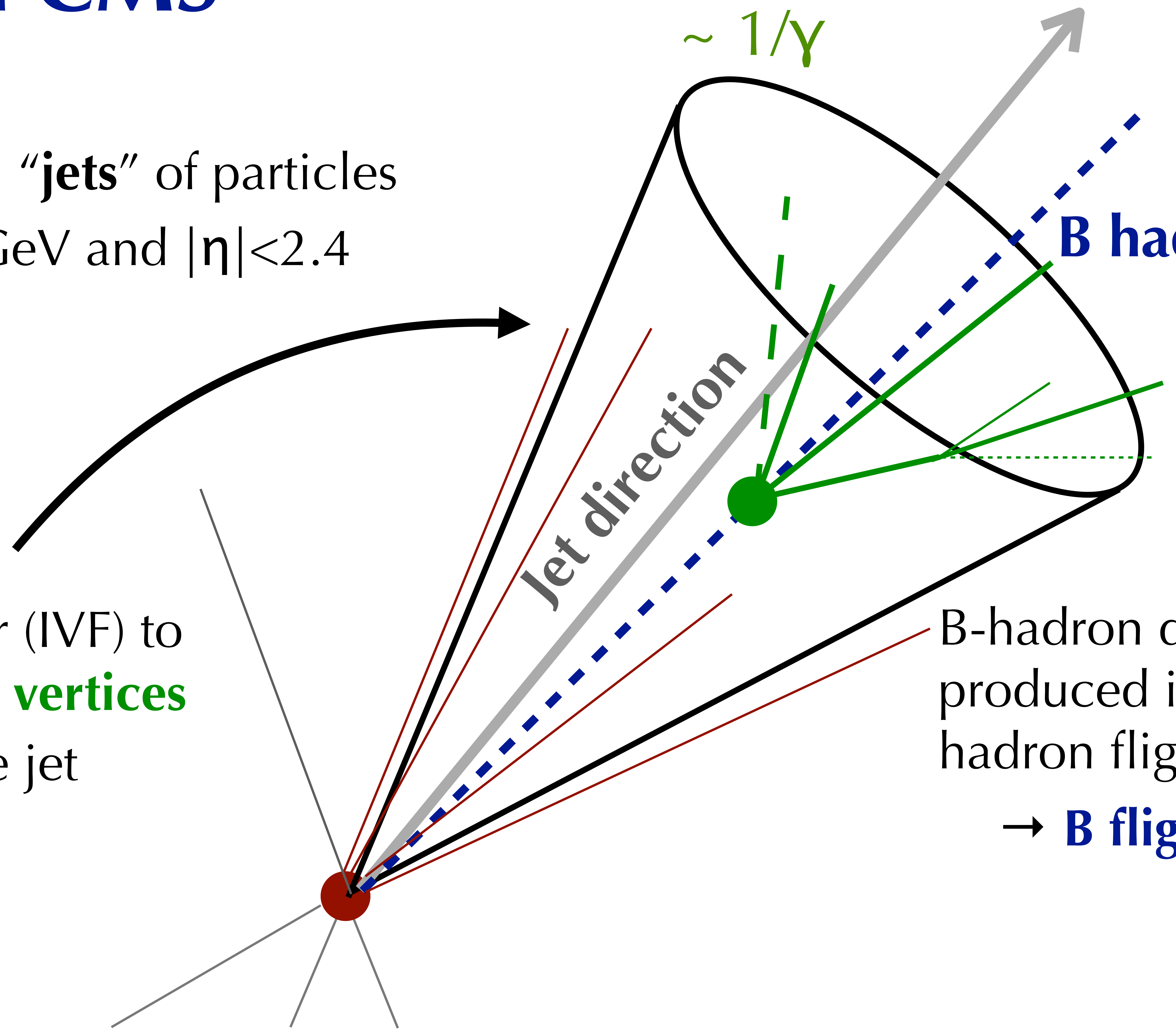


- Higgs discovery and H → $b\bar{b}$ status at LHC
  - challenges of H → $b\bar{b}$ at LHC
- **b-tagging**
  - dedicated strategy for boosted H → $b\bar{b}$
- H → $b\bar{b}$ tagging
- Inclusive search for boosted H → $b\bar{b}$

b tagging in CMS

b quarks hadronize in "**jets**" of particles

- Jets with p_T> 30 GeV and $|\eta|$<2.4



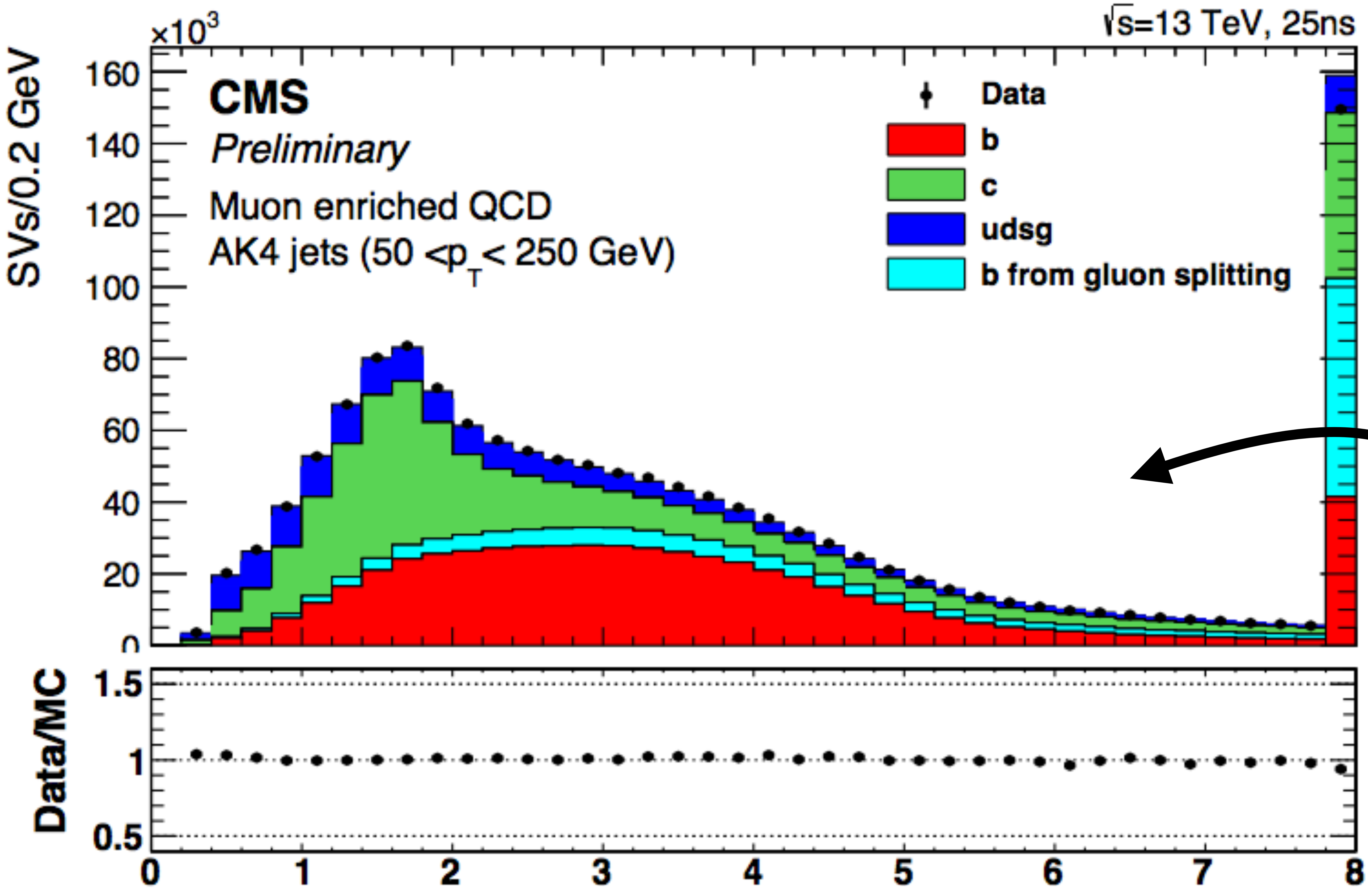
Inclusive Vertex Finder (IVF) to reconstruct **secondary vertices**

- Independent of the jet direction

B-hadron decay tracks are mostly produced in a cone in the B hadron flight direction

→ **B flight ~ jet direction**

b tagging in CMS



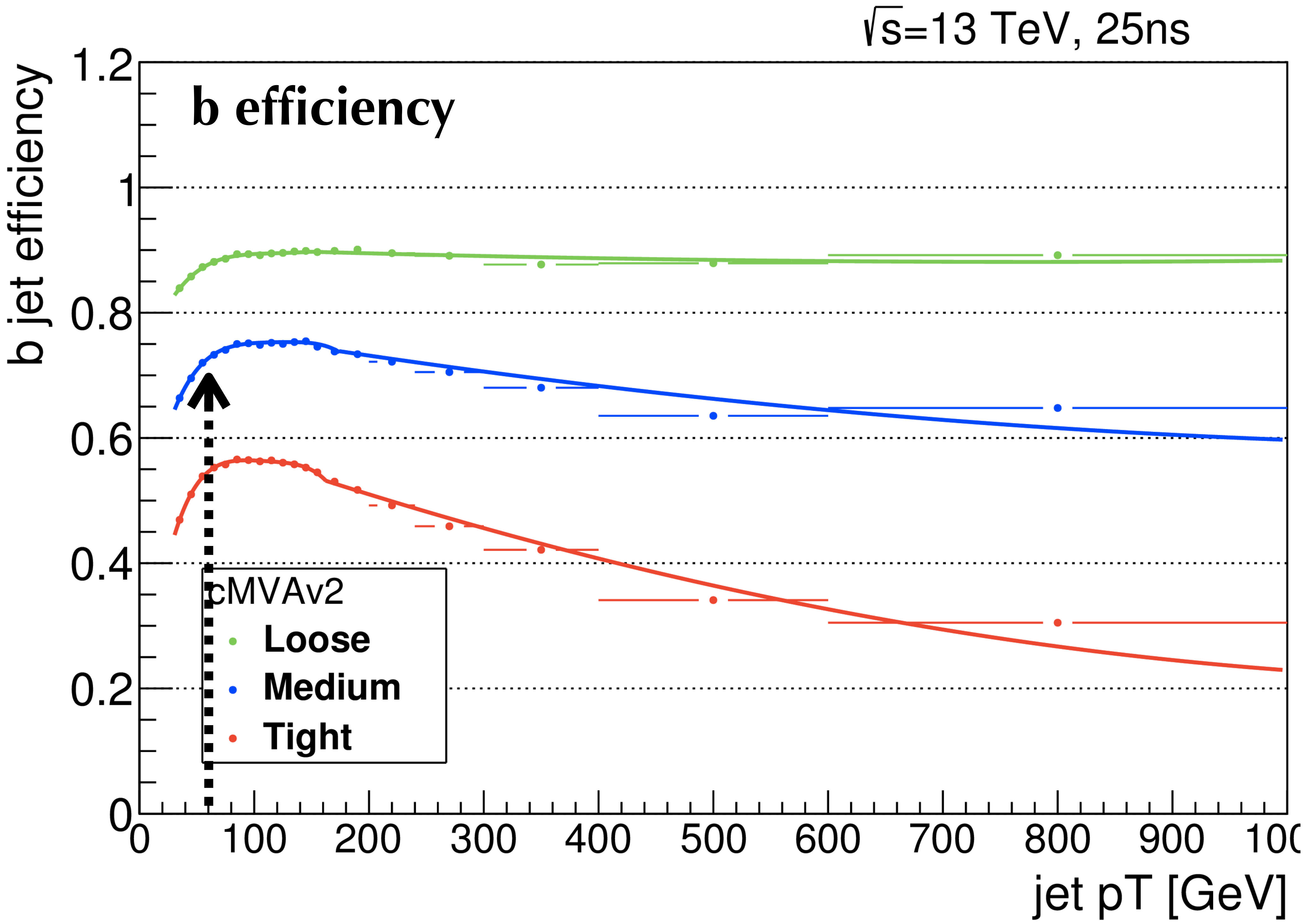
secondary vertex mass (GeV)

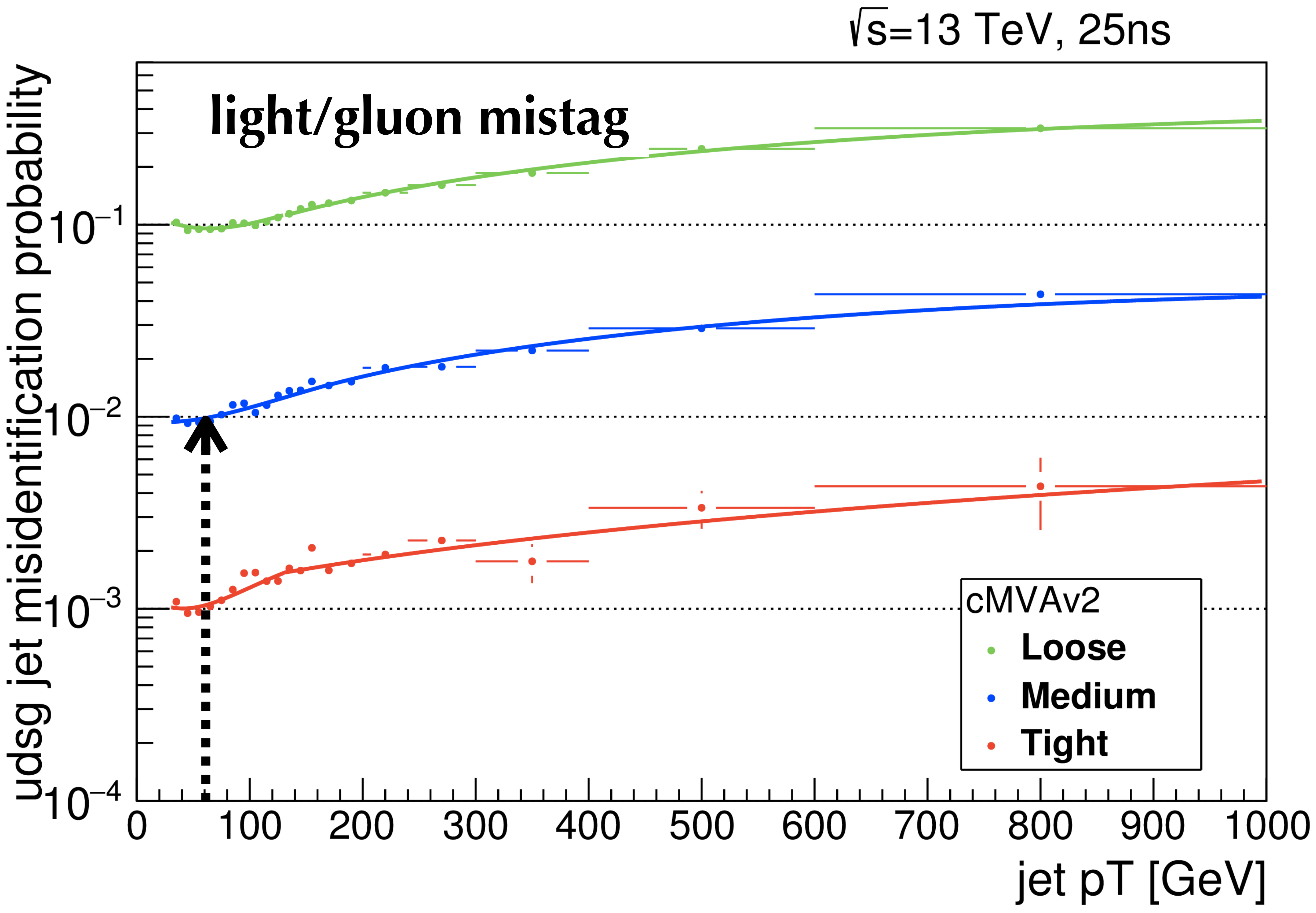
b-tagging algorithms combine with a **multivariate approach** the information from:

- **impact parameter** significance of charged-particle tracks
- the presence and properties of reconstructed **secondary vertices**
- the presence of a **lepton** in the jet and its properties

Performance of b-tagging in CMS

CMS-PAS-BTV-15-001





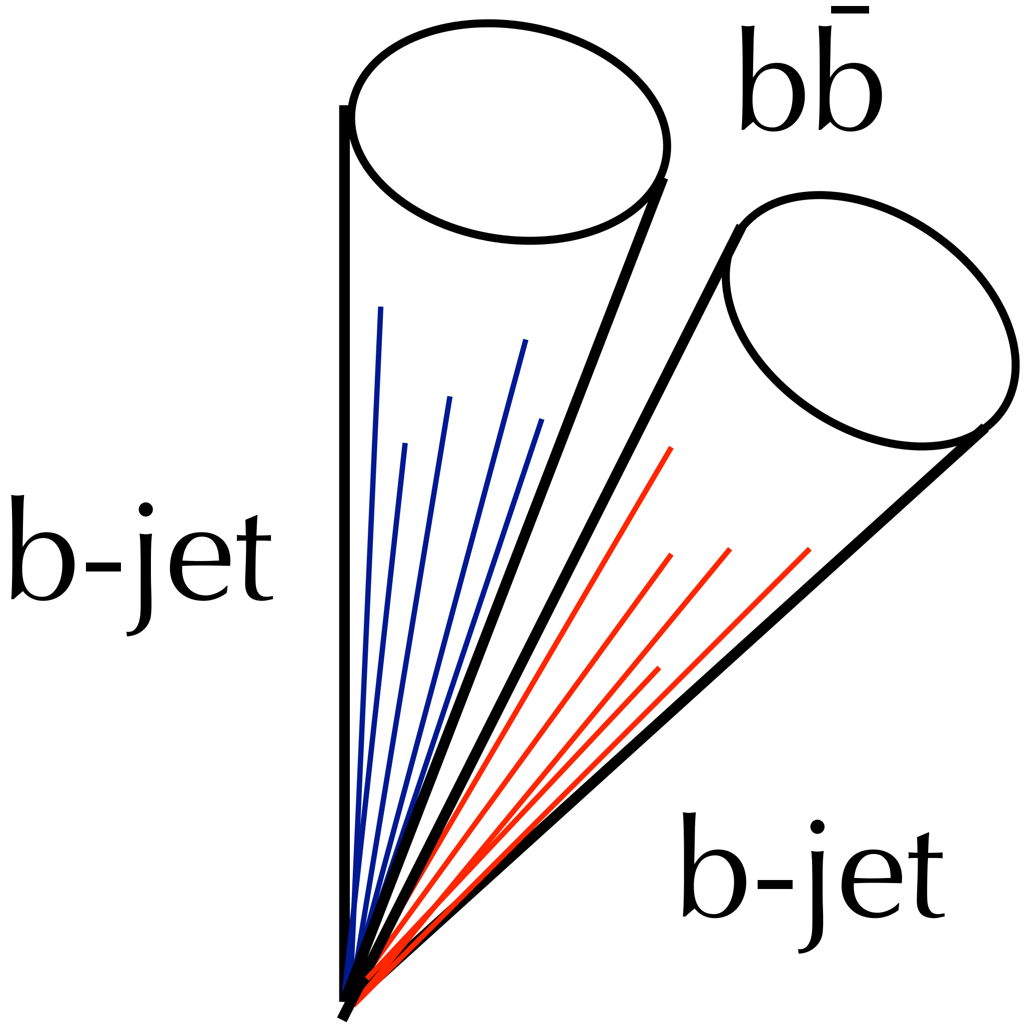
optimal working point for a **H to $b\bar{b}$ search** search is 70% b efficiency and 1% mistag probability

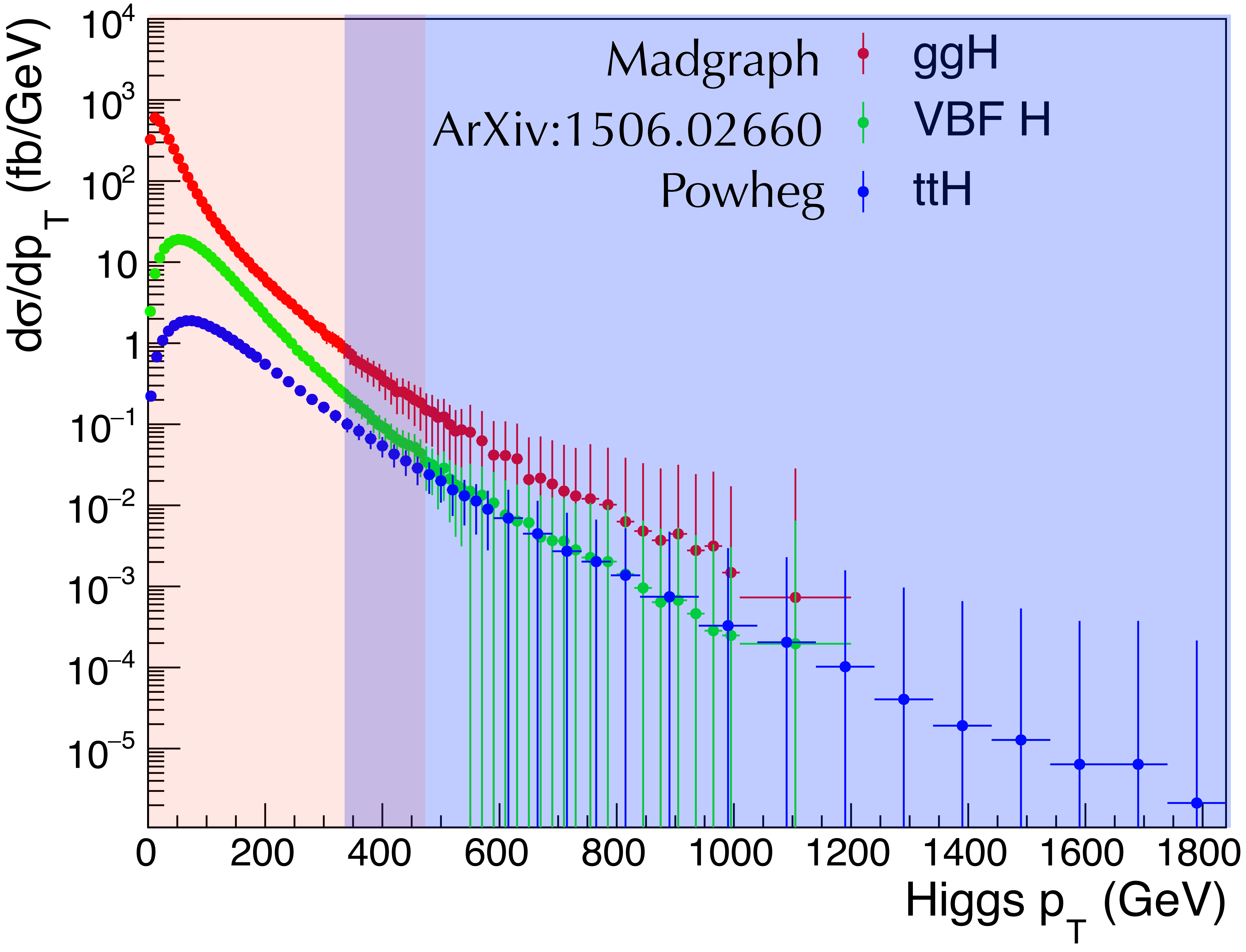
boosted $H(b\bar{b})$

$dR(b\bar{b}) \sim 2m_H/p_T$

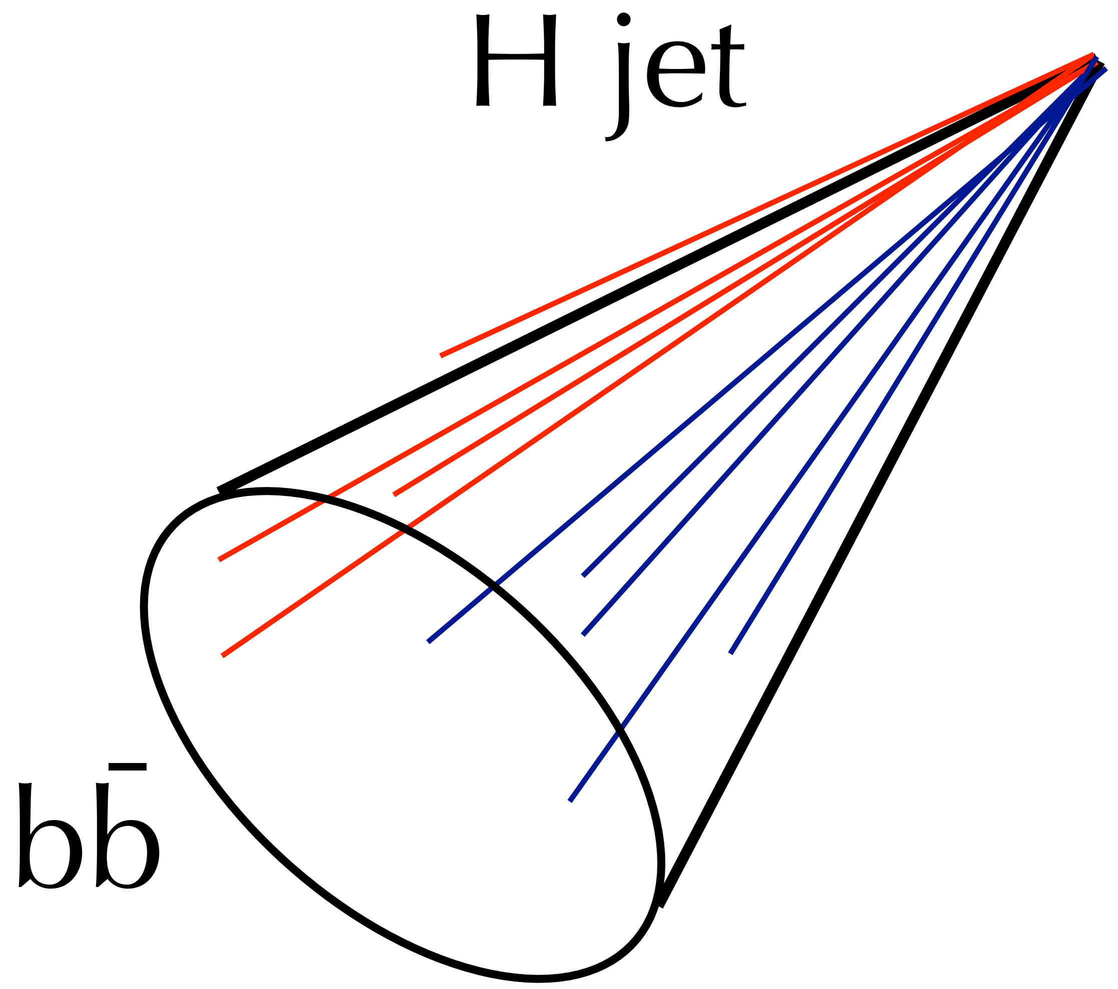
300 400 500 600 H p_T (GeV)

two-separate b-jets (R = 0.4)

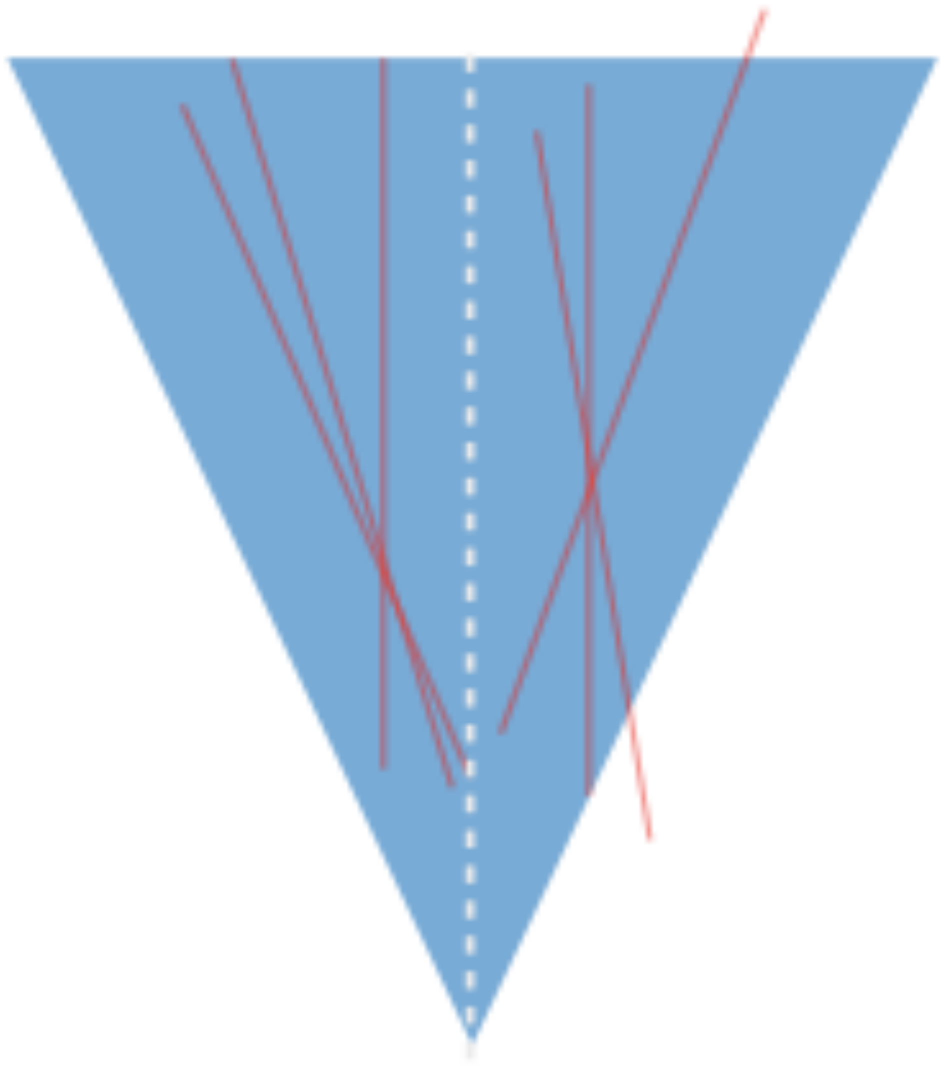




one single large-cone (fat) jet (R = 0.8)

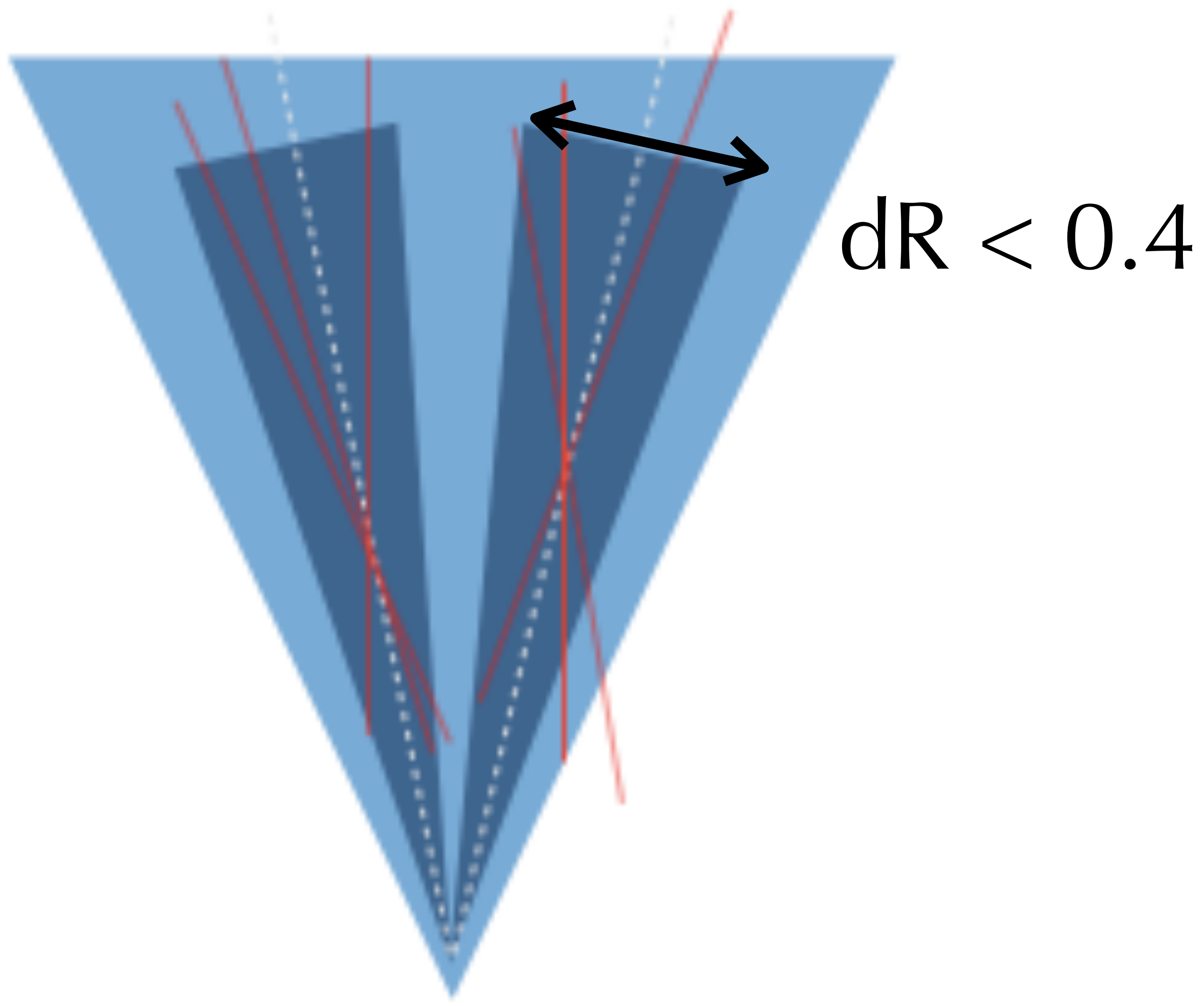


Multiple approaches



fat-jet b-tagging

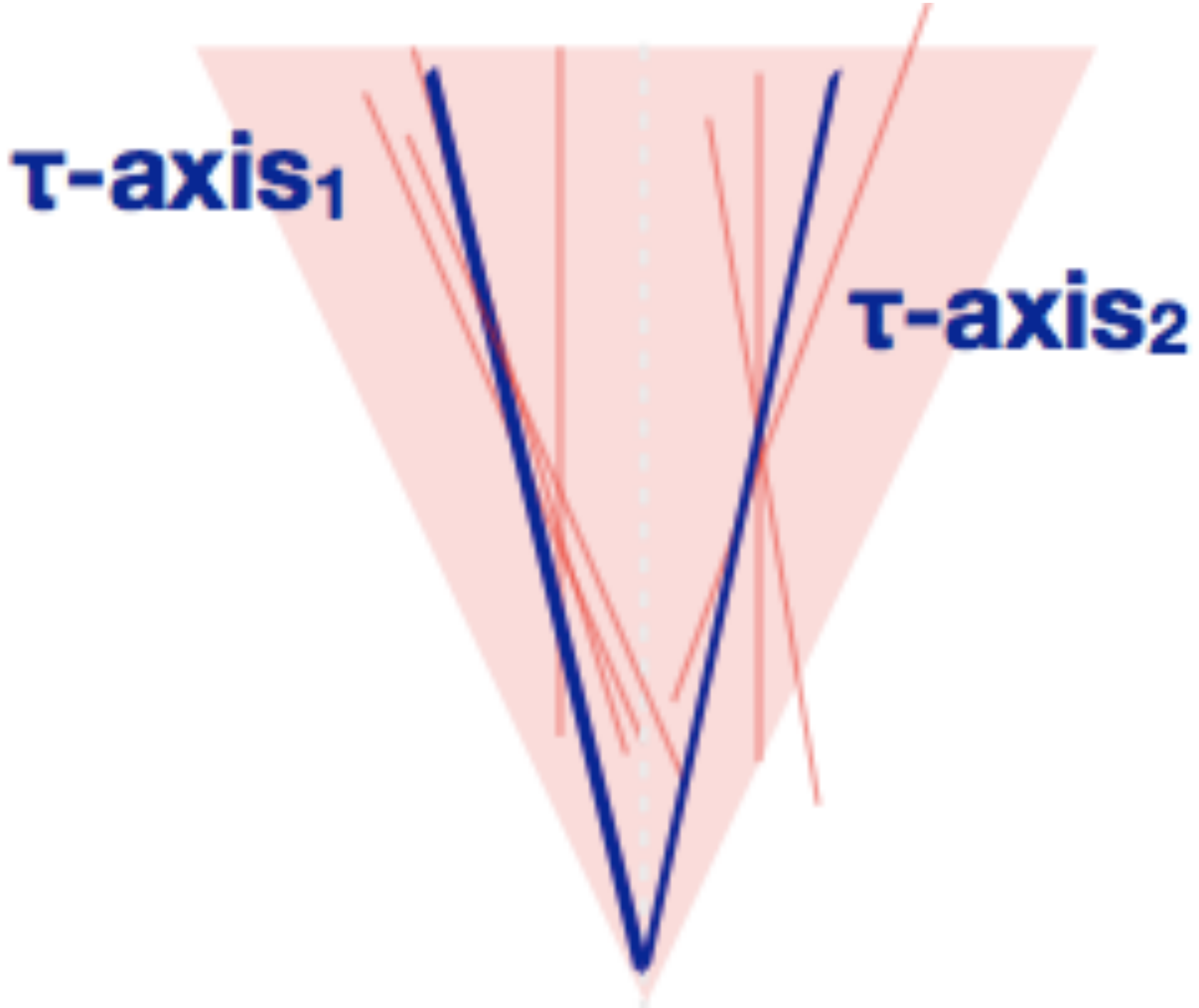
- Based on the standard b-tagging algorithm
- Not designed for tagging two b's in the same jet



sub-jet b-tagging

- Defines sub-jets
- Standard b-tagging algorithm applied to each subjet

NEW



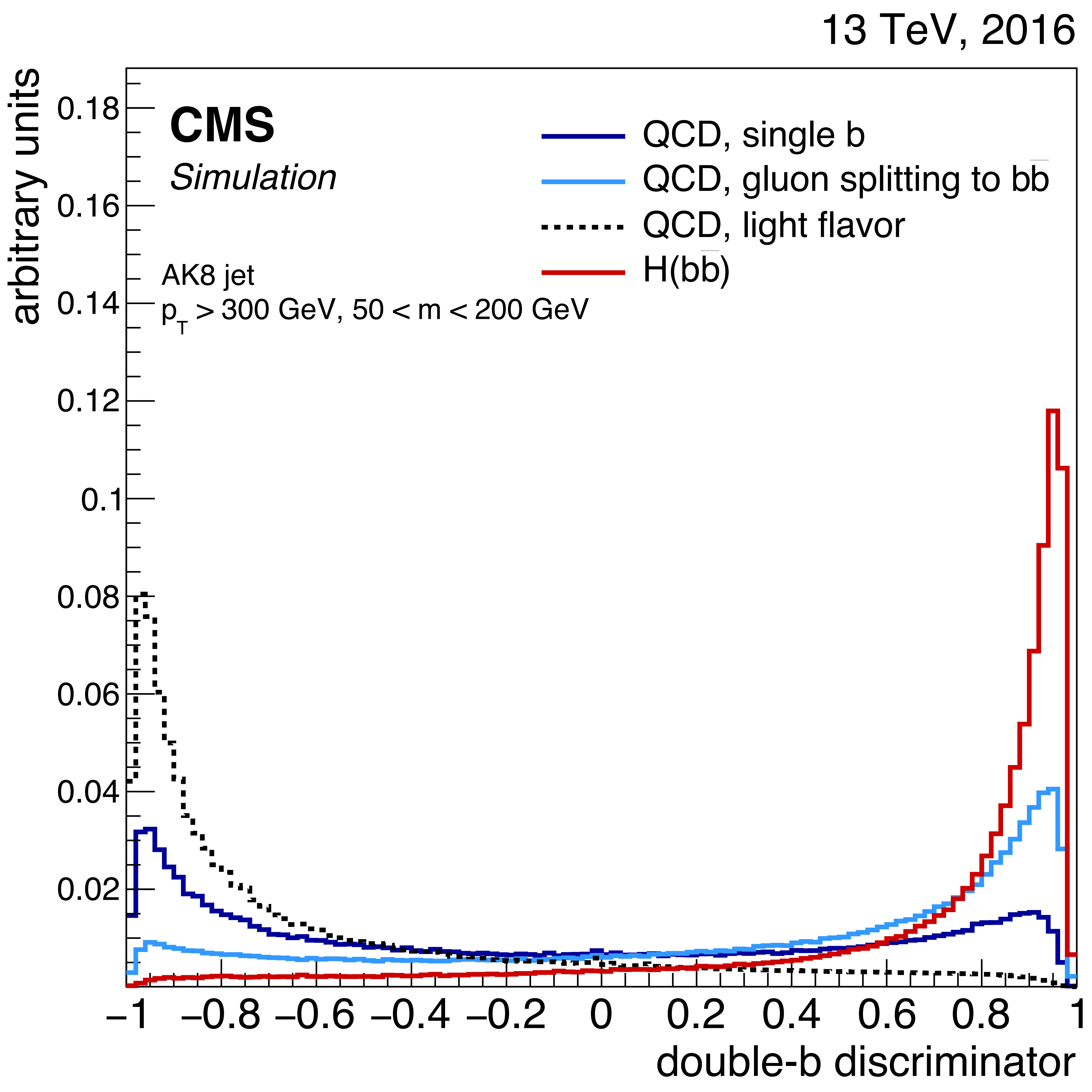
double-b tagger

- Identifies the two B hadron decay chains from b and $\bar{b}$ within the same fat jet.
- It does not define sub-jet but uses N-jettiness axes

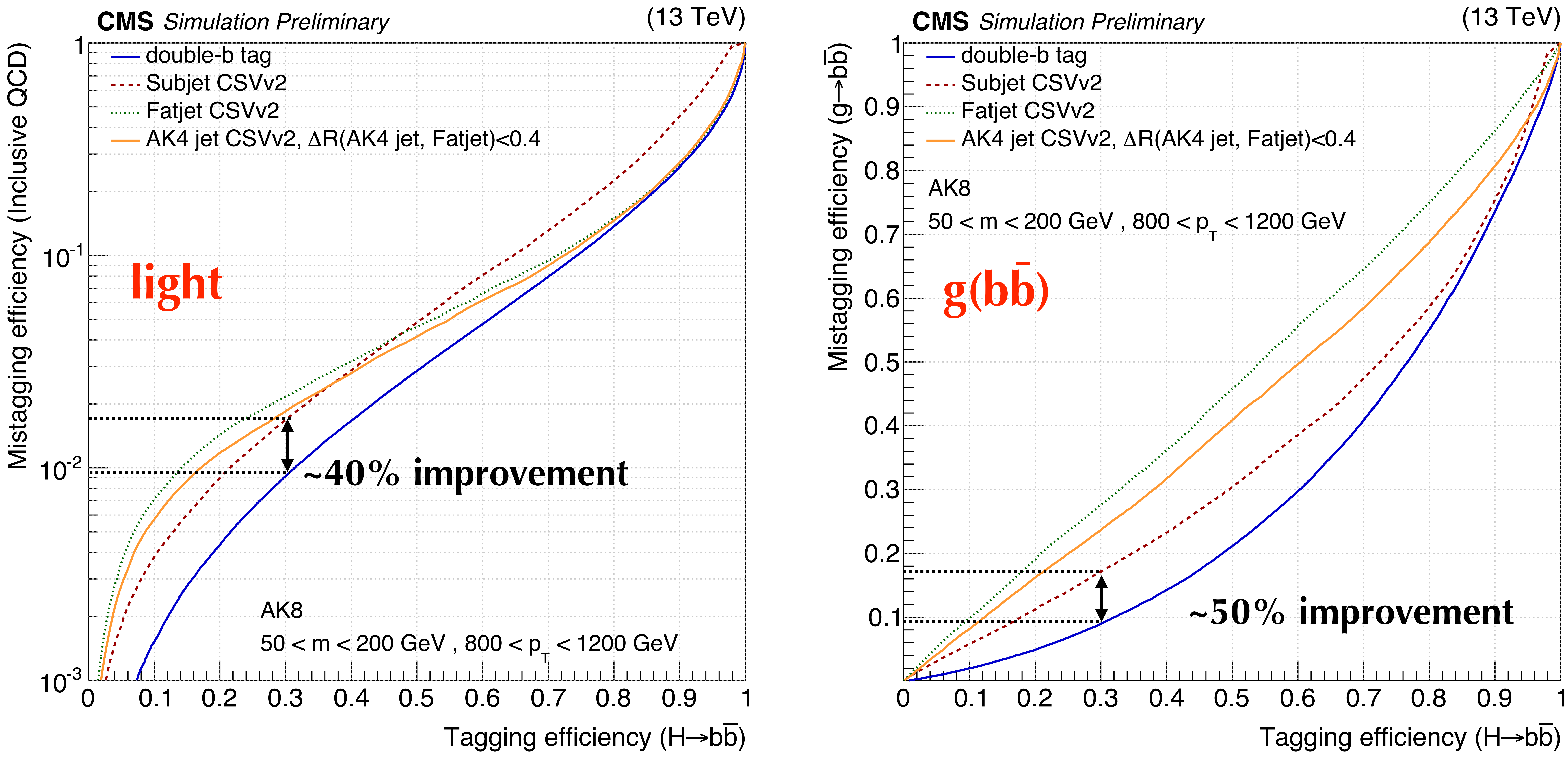
*τ-axes are Nsubjettiness axis

double-b tagger

- Combines **tracking and vertexing information** with a multivariate approach
  - 27 observables are used
- It **targets the $b\bar{b}$** signal aiming to be:
  - ***mass independent***
  - ***p_T independent***
  - **training strategy** is designed to cover a very wide p_T range
  - inputs are chosen to avoid p_T correlation
    - *no dR-like variables, no substructure info*

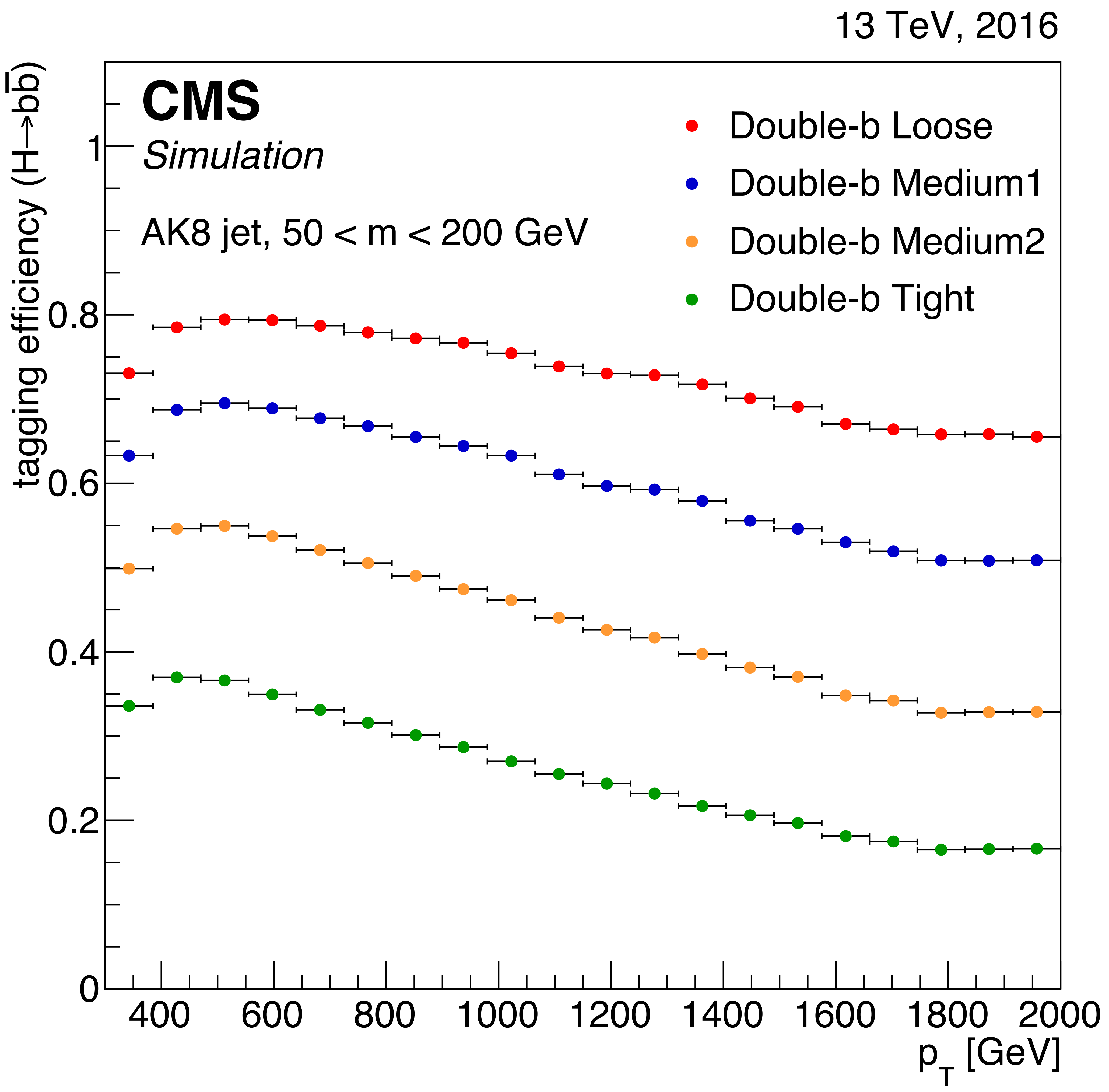


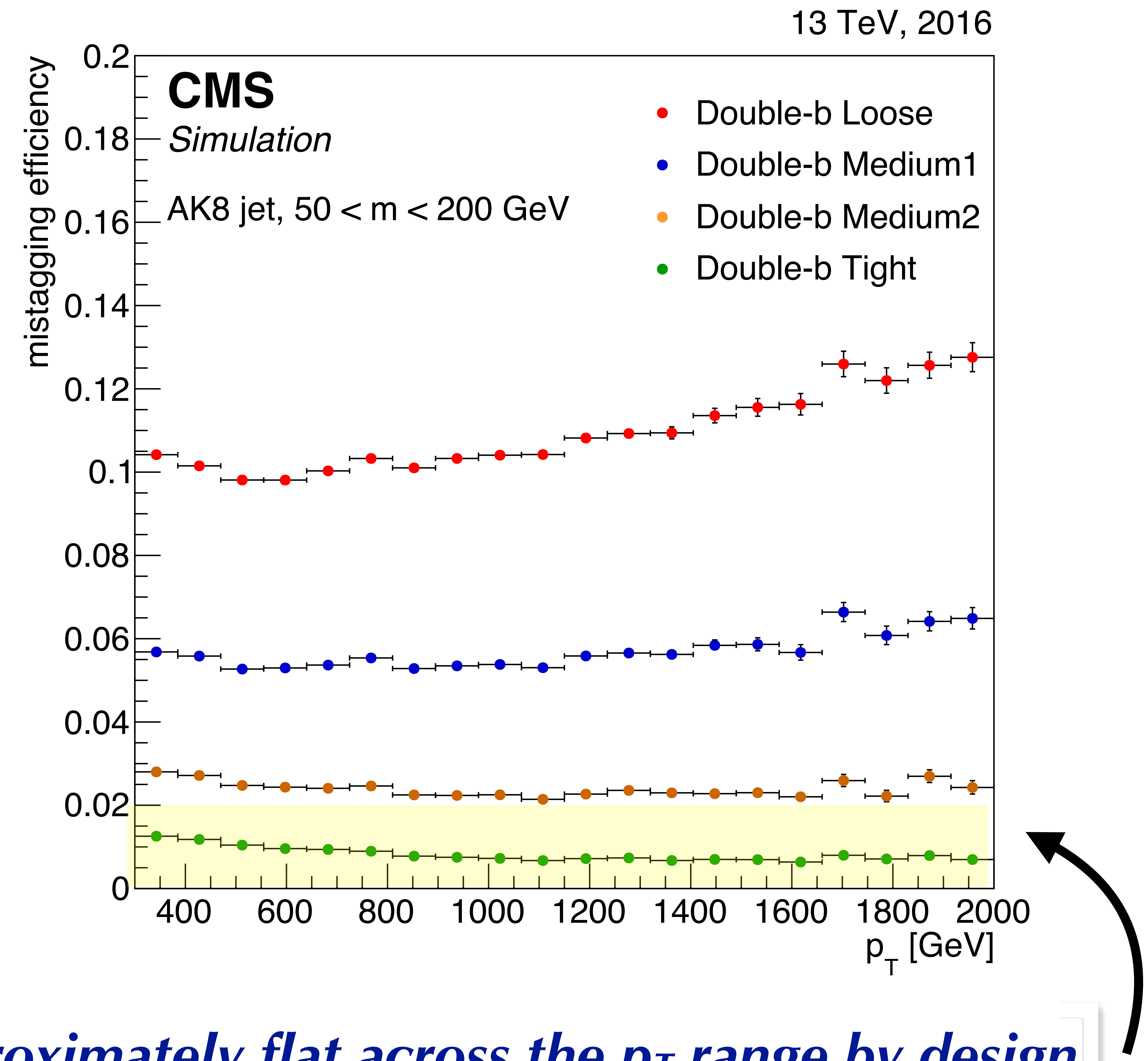
Efficiency vs. Mistag rate



Mistag is reduced by more than 40% at 30% signal efficiency (~ tight working point)

Performance

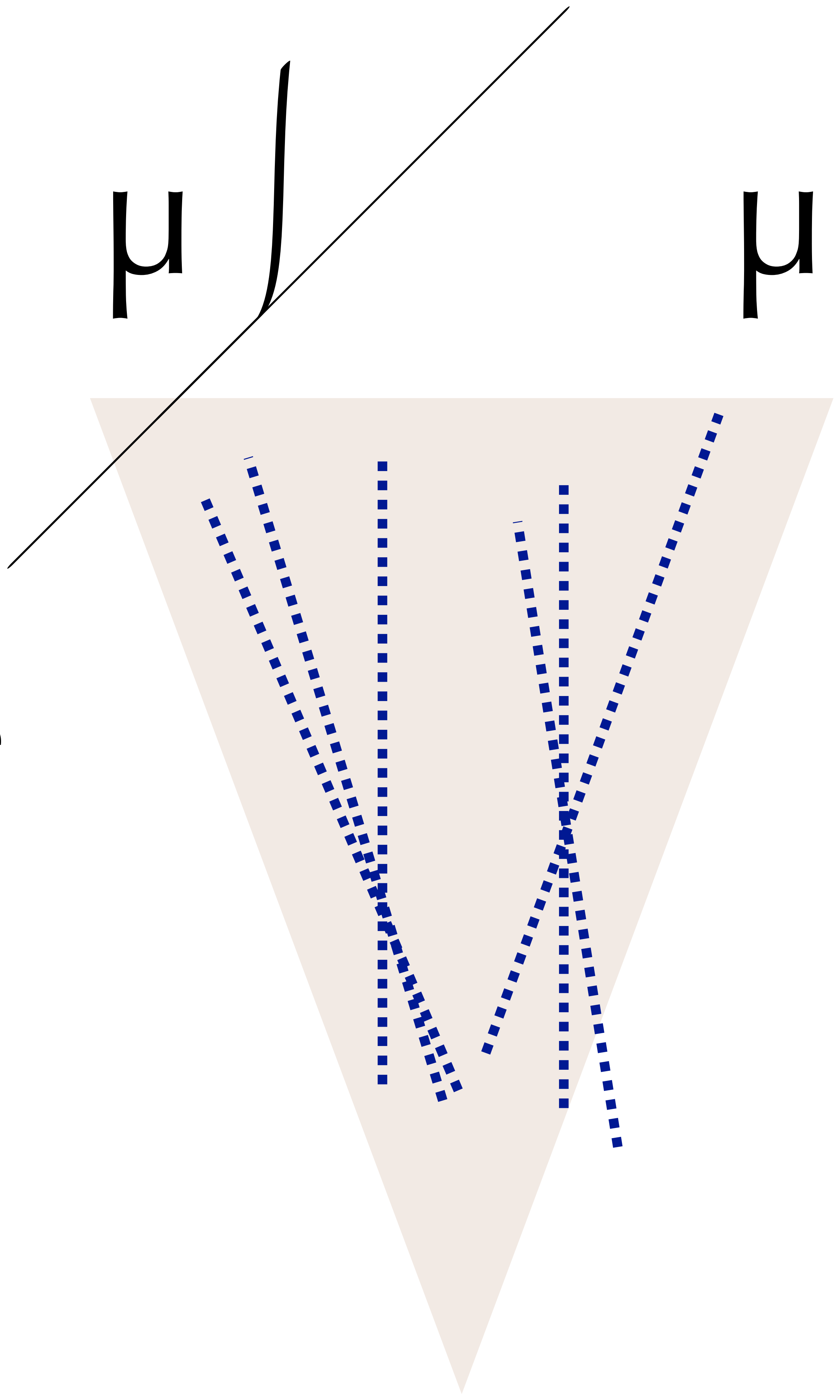




The mistag rate is approximately flat across the p_T range by design

Critical point for searches (background estimate)

Efficiency measurement in data



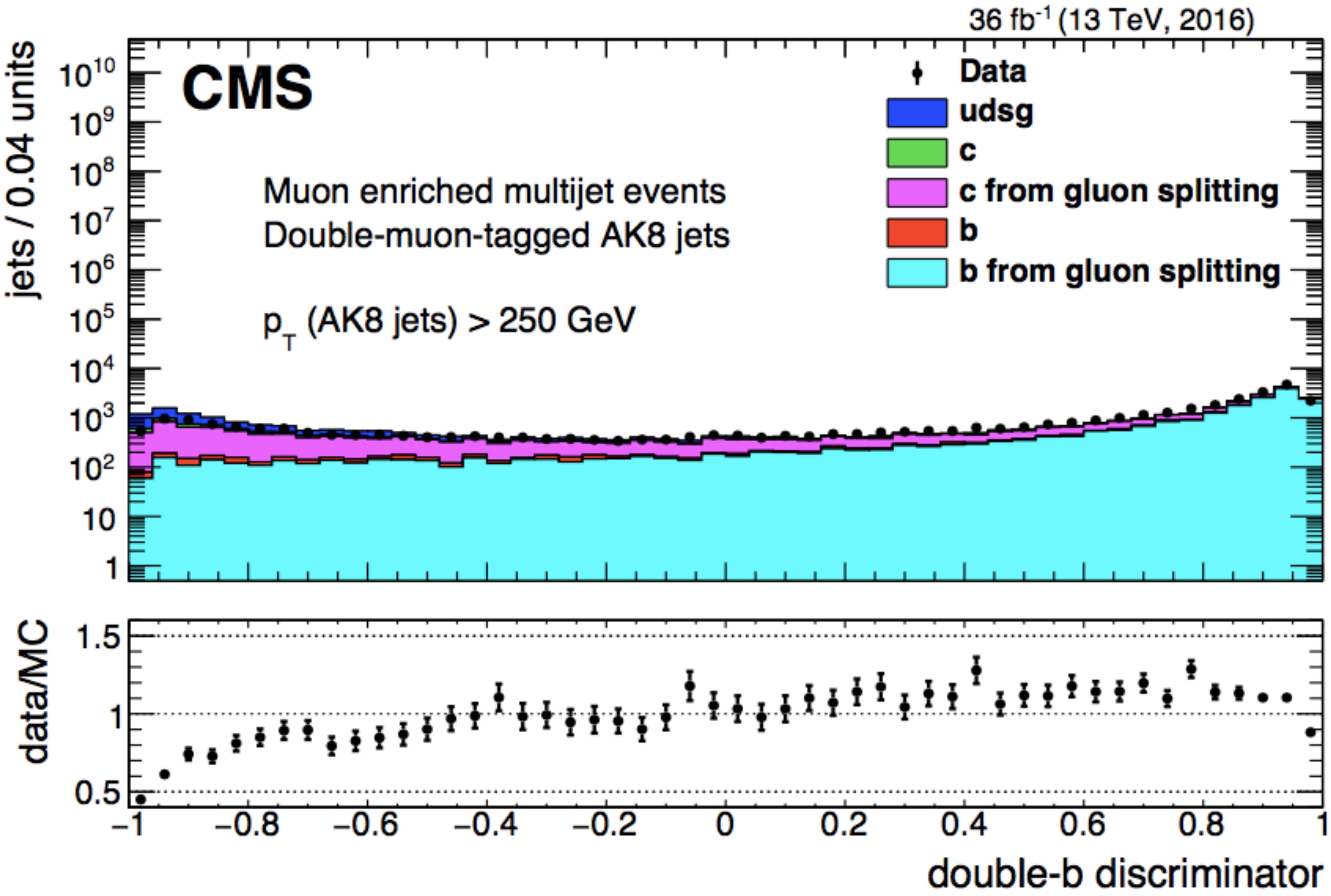
- Since there is no H/Z($b\bar{b}$) signal (yet!) we use:
  - **g($b\bar{b}$) jets as a proxy** to measure the signal efficiency
- Jet selection has been designed to ensure jets are signal-like
  - High AK8 p_T jet (p_T > 250 GeV)
  - **double-muon** tagged jets (muon with p_T > 7 GeV)
  - **mass cut** (>50 GeV)

Z($b\bar{b}$) by the end of the talk

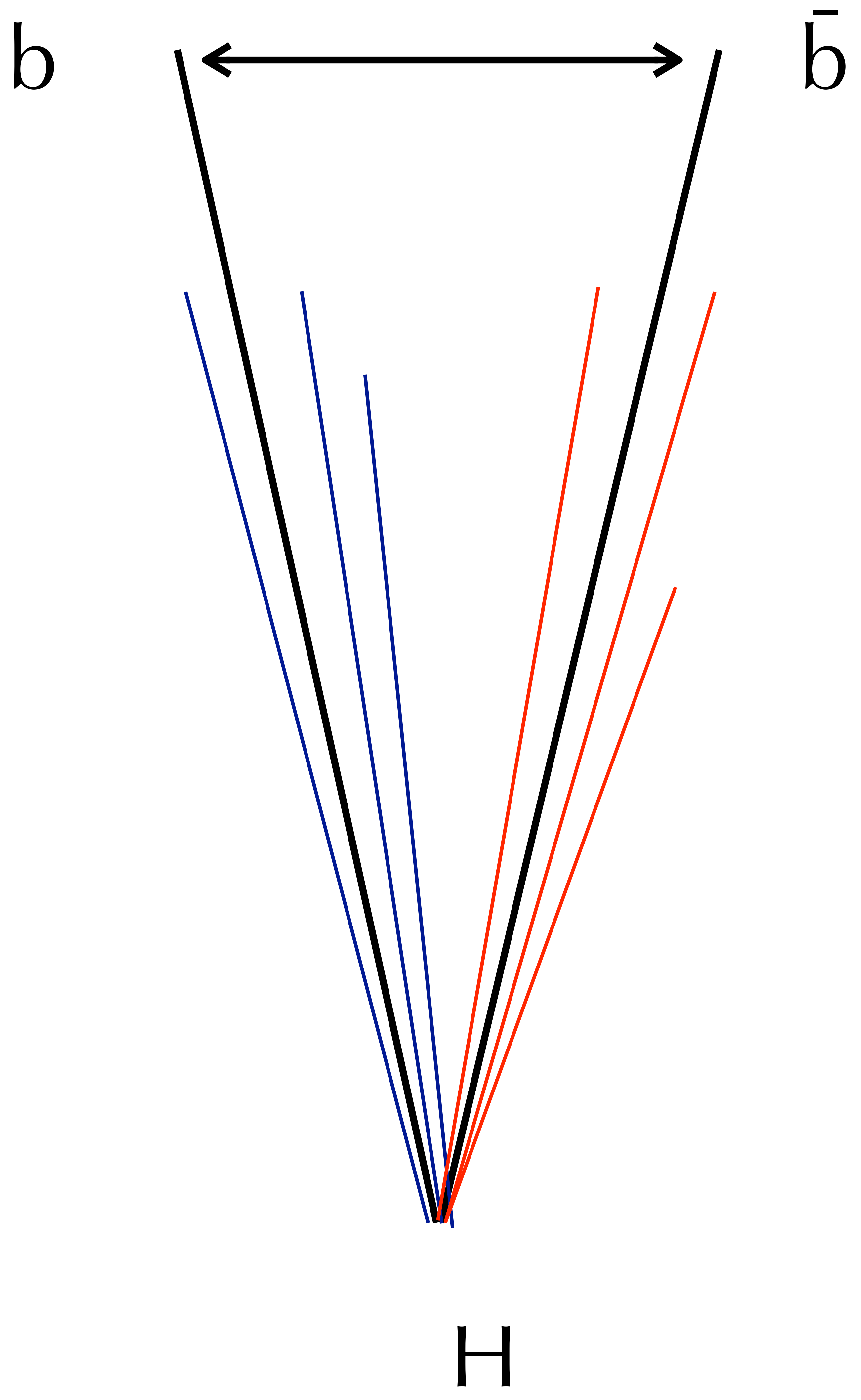
Efficiency measurement in data

CMS-PAS-BTV-16-002

Associated **uncertainty** varies from **3 to 5%** depending on the different tagging efficiency



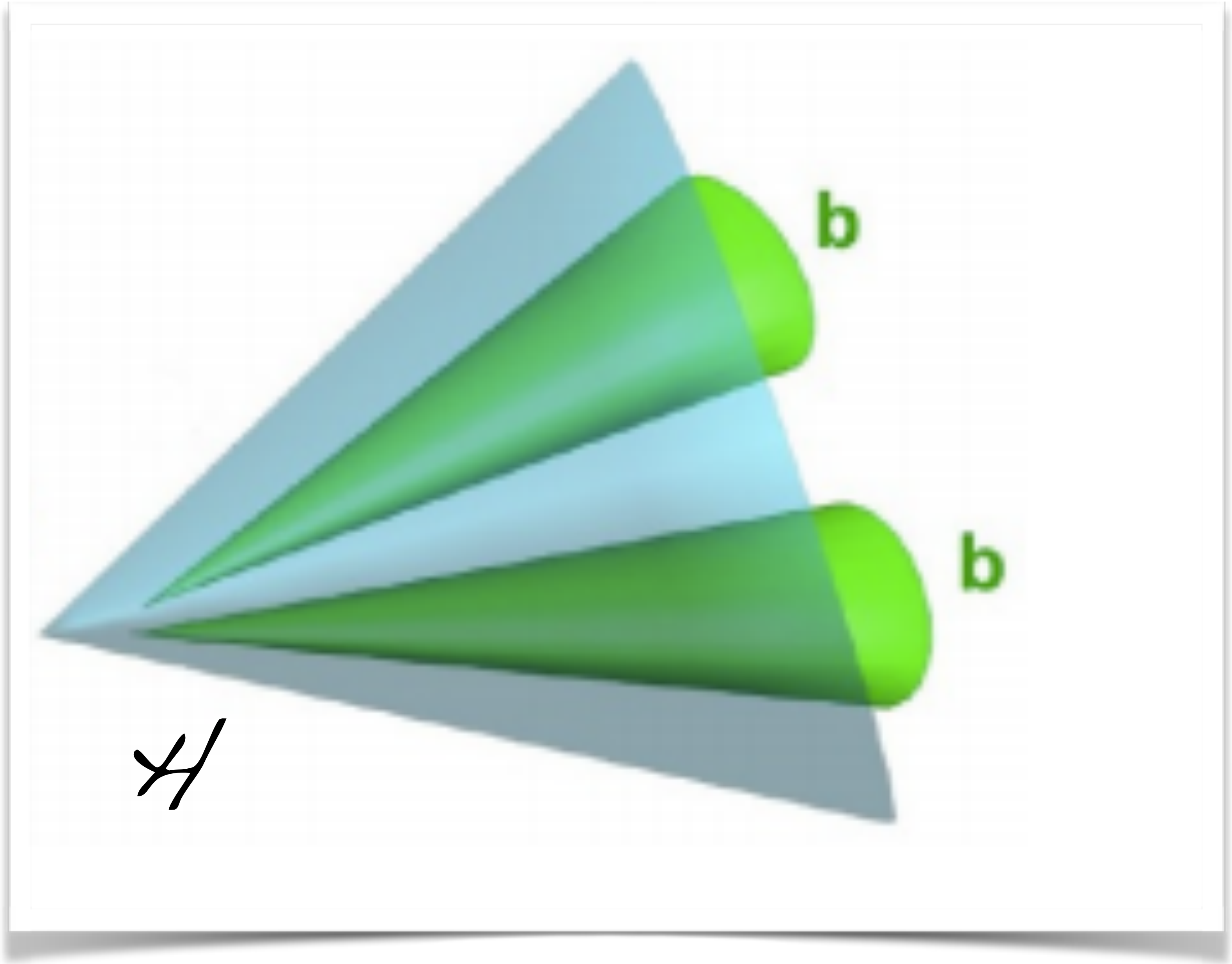
- Higgs discovery and $H \rightarrow b\bar{b}$ status at LHC
    - challenges of $H \rightarrow b\bar{b}$ at LHC
- b-tagging
    - dedicated strategy for boosted $H \rightarrow b\bar{b}$
- **$H \rightarrow b\bar{b}$ tagging**
- Inclusive search for boosted $H \rightarrow b\bar{b}$

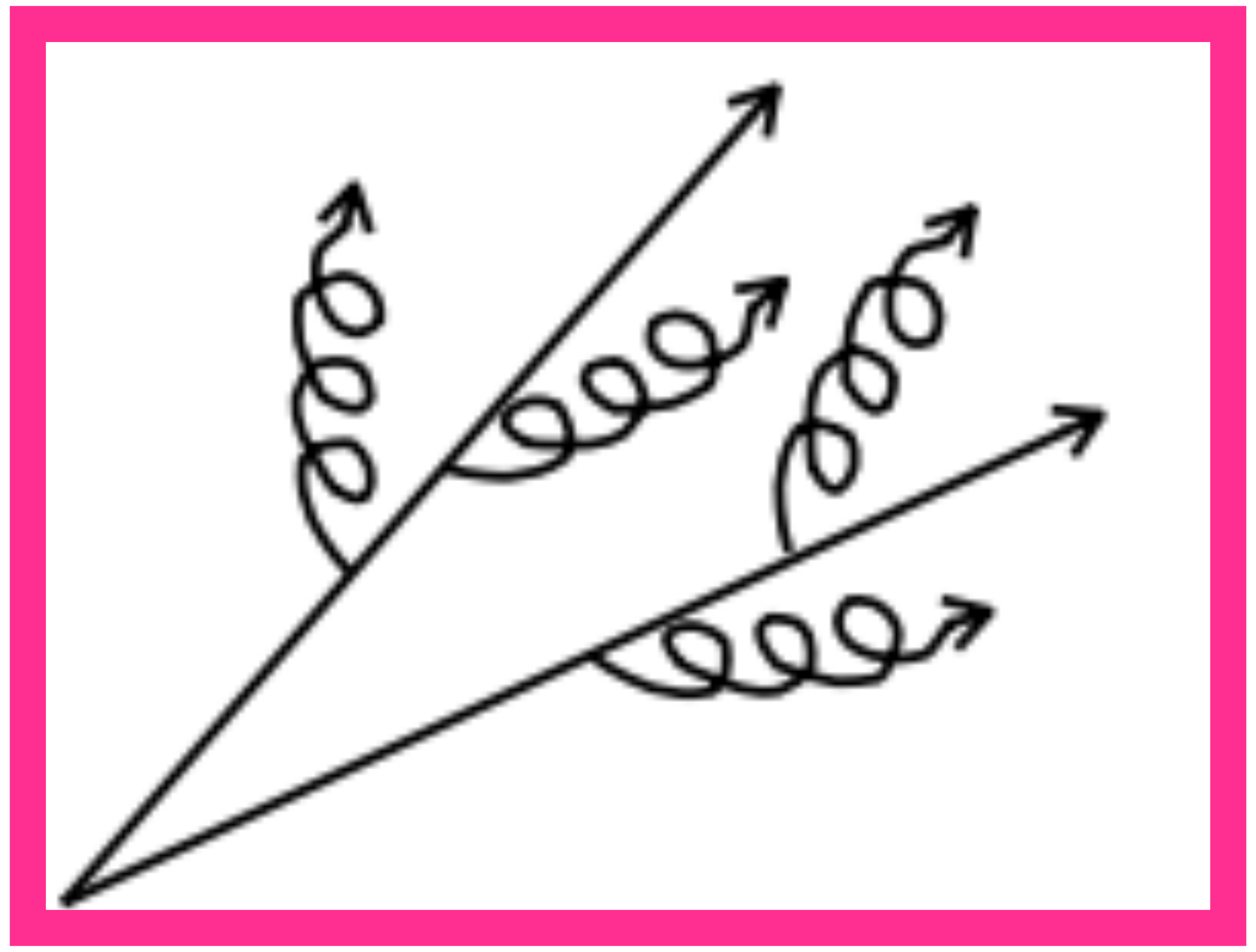


H($b\bar{b}$) tagging

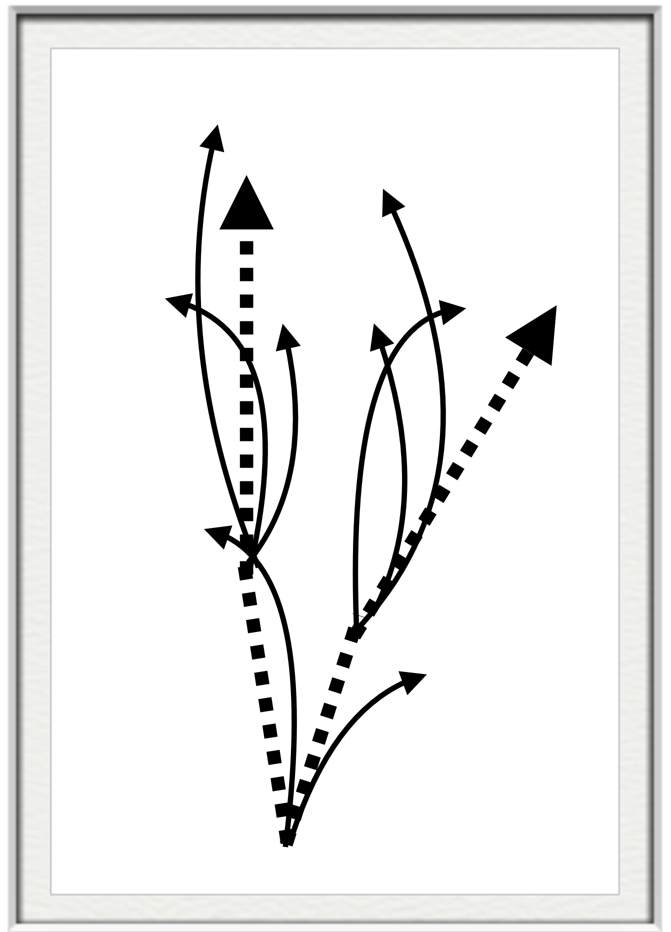
The boosted H($b\bar{b}$) signal is identified as large cone size jets:

- anti-k_T algorithm with R=0.8
- **PUPPI** (PileUp Per Particle Id) is used to mitigate **pile up effects**

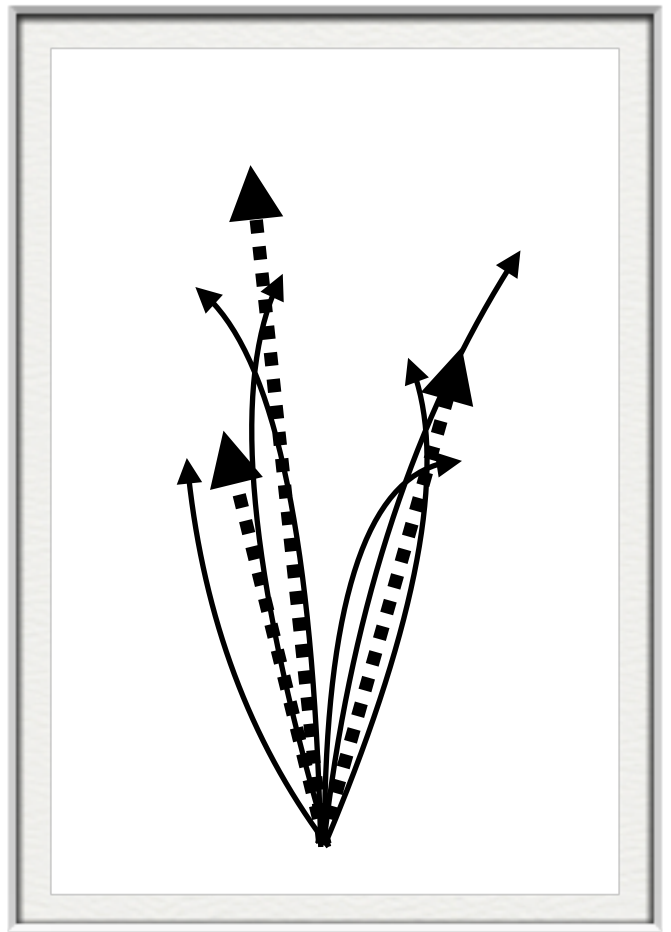




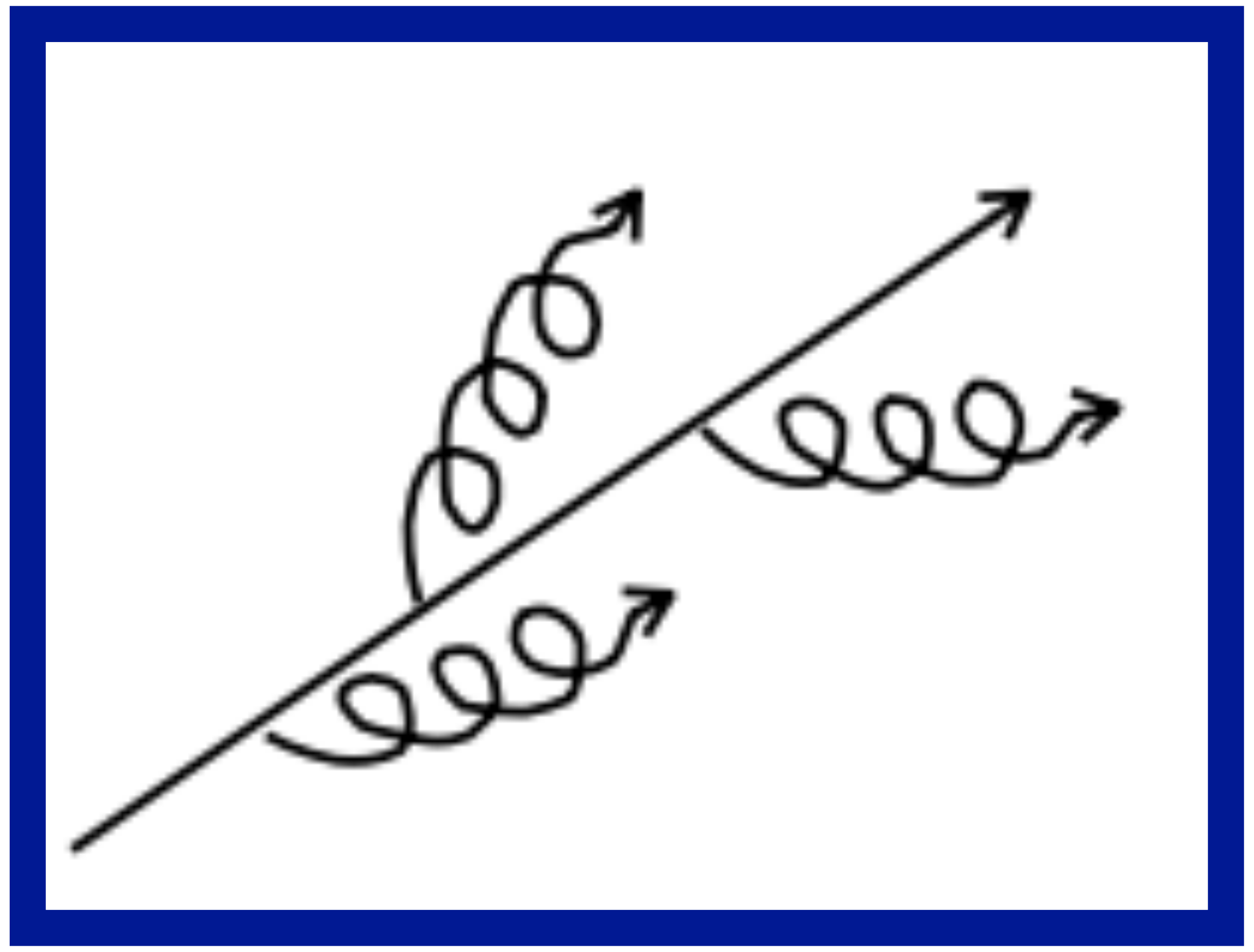
signal



H/Z(bb)



W/Z(qq)



background q/g

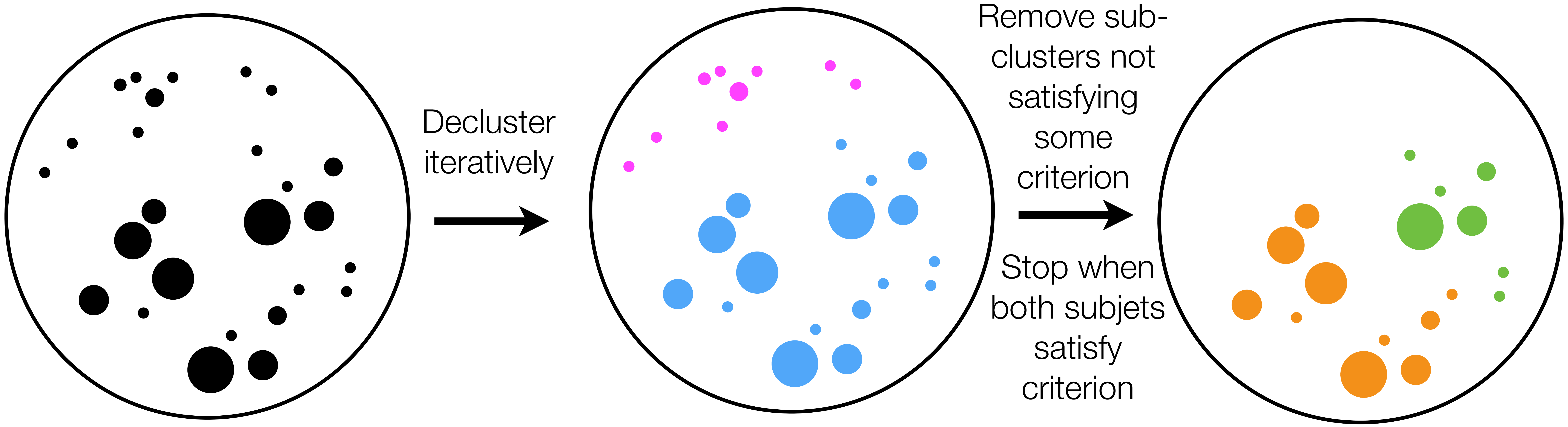
Our tools:

- **b-tagging** to reconstruct the two B hadrons from the b and $\bar{b}$ within the same fat jet
- jet **mass** compatibility with the Higgs
- the composite nature of the jet using **substructure**

ArXiv:1307.0007
ArXiv:1402.2657

Jet mass

- Provides good separation between W/Z/H-jets from q/g jets
- **Grooming** removes soft and wide-angle radiation (**soft drop/modified mass soft drop**)



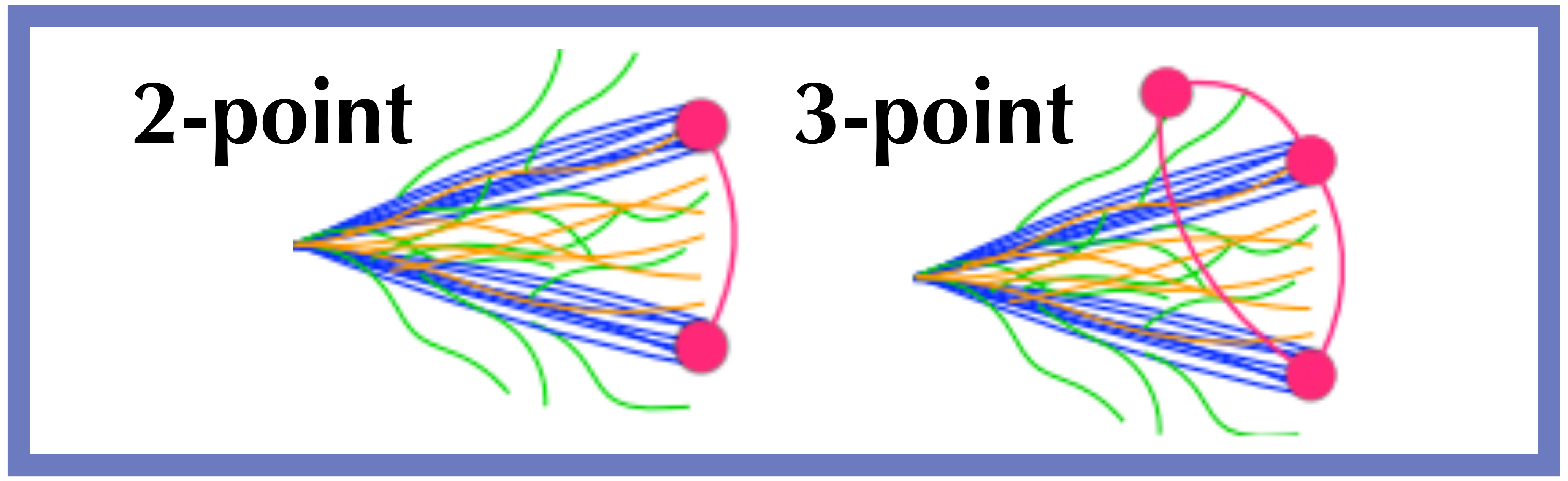
$$\frac{\min(p_{T1}, p_{T2})}{p_{T1} + p_{T2}} > z_{\text{cut}} \left(\frac{\Delta R_{12}}{R_0}\right)^{\beta}$$

$\beta = 0,\ z_{cut} = 0.1$

Jet Substructure

- Measures the degree to which a jet can be considered as composed of N prongs
- **Energy correlation functions** are sensitive to N-point correlations in a jet
  - A 2-pronged jet will have $e_3 < e_2$

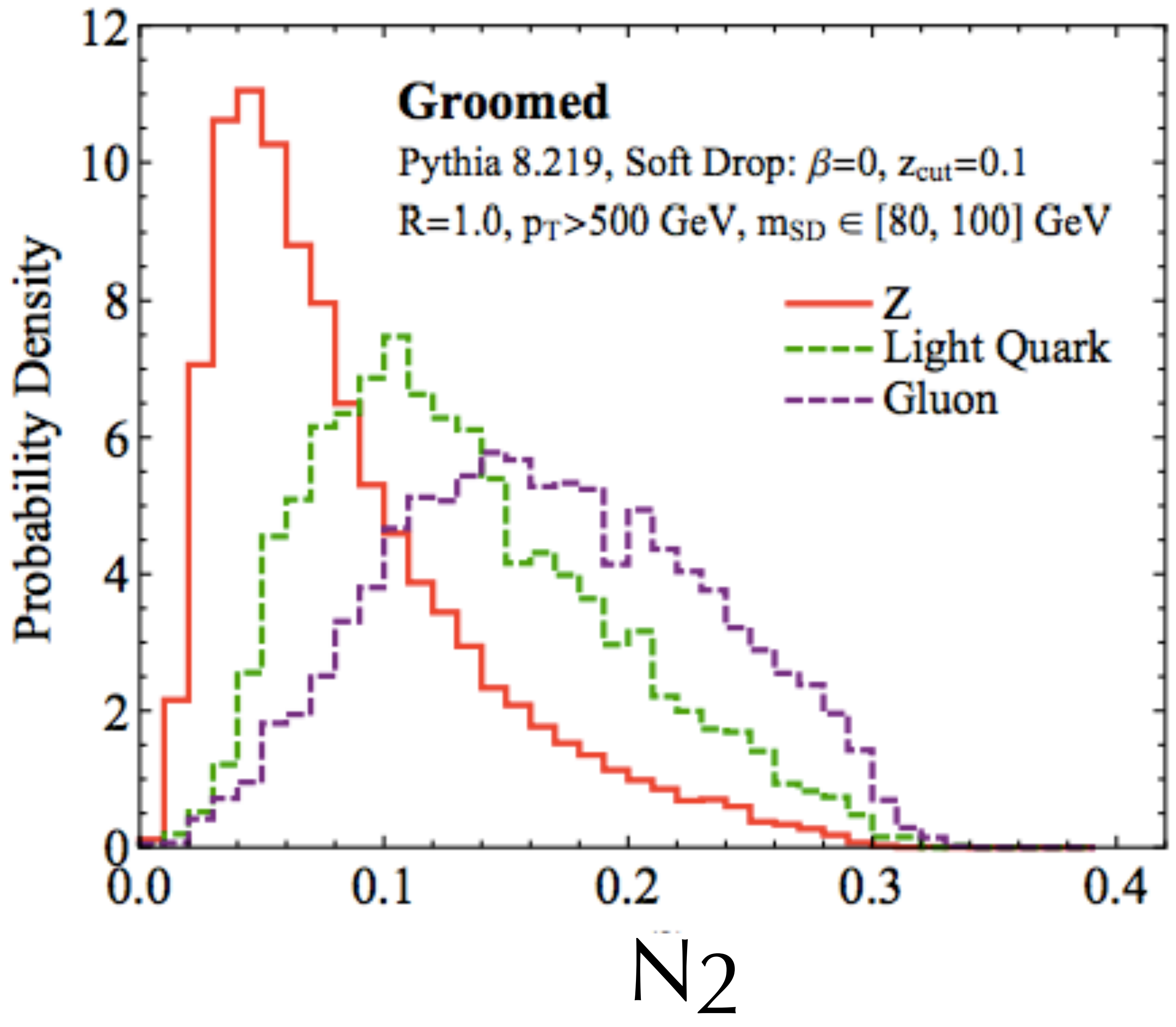
$$N_2^\beta = \frac{{}_2e_3^\beta}{({}_1e_2^\beta)^2}$$



$${}_1e_2^\beta = \sum_{1\ \leq i<j\ \leq n_j} z_i z_j \Delta R_{ij}^\beta$$

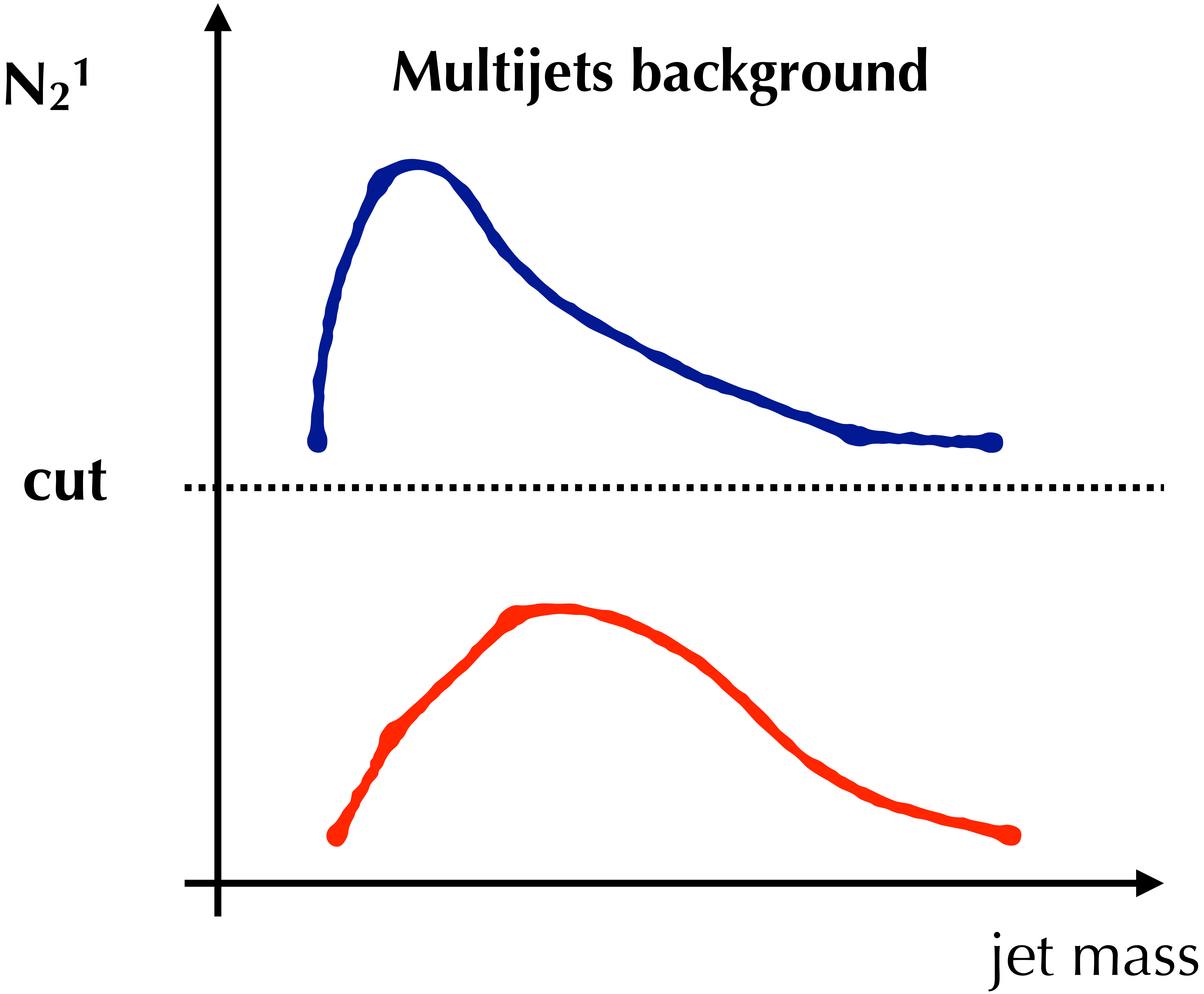
$${}_2e_3^\beta = \sum_{1\ \leq i<j<k \leq n_j} z_i z_j z_k \ \min\left\{\Delta R_{ij}^\beta \Delta R_{ik}^\beta, \Delta R_{ij}^\beta \Delta R_{jk}^\beta, \Delta R_{ik}^\beta \Delta R_{jk}^\beta\right\}$$

β=1



Jet Substructure

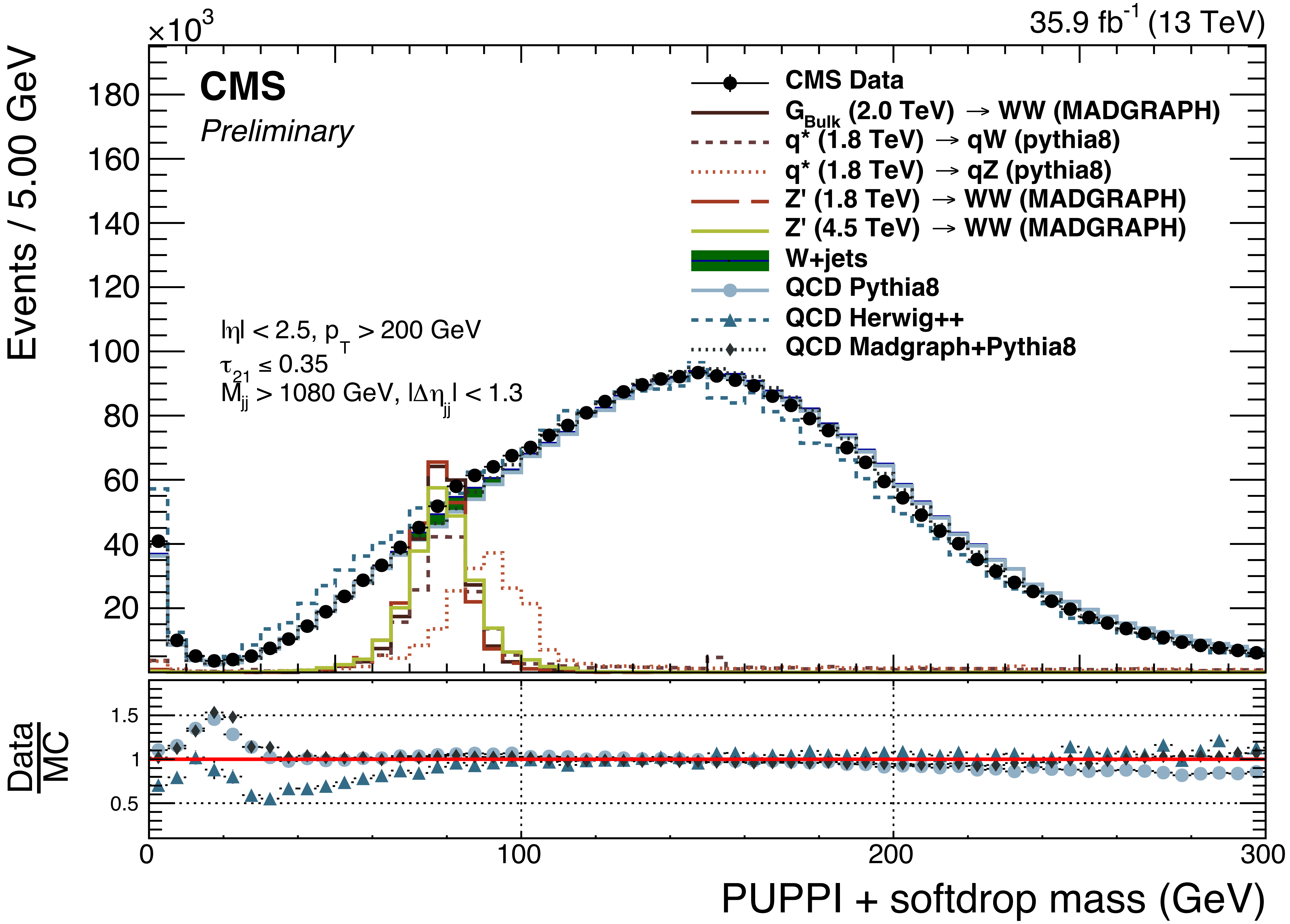
ArXiv:1603.00027
CMS-PAS-B2G-17-001
CMS-PAS-EXO-17-001



$N^1{}_2$ sculpts jet mass distribution

Jet Substructure

ArXiv:1603.00027
CMS-PAS-B2G-17-001
CMS-PAS-EXO-17-001



N^1_2 sculpts jet mass distribution

Jet Substructure

ArXiv:1603.00027
CMS-PAS-B2G-17-001
CMS-PAS-EXO-17-001

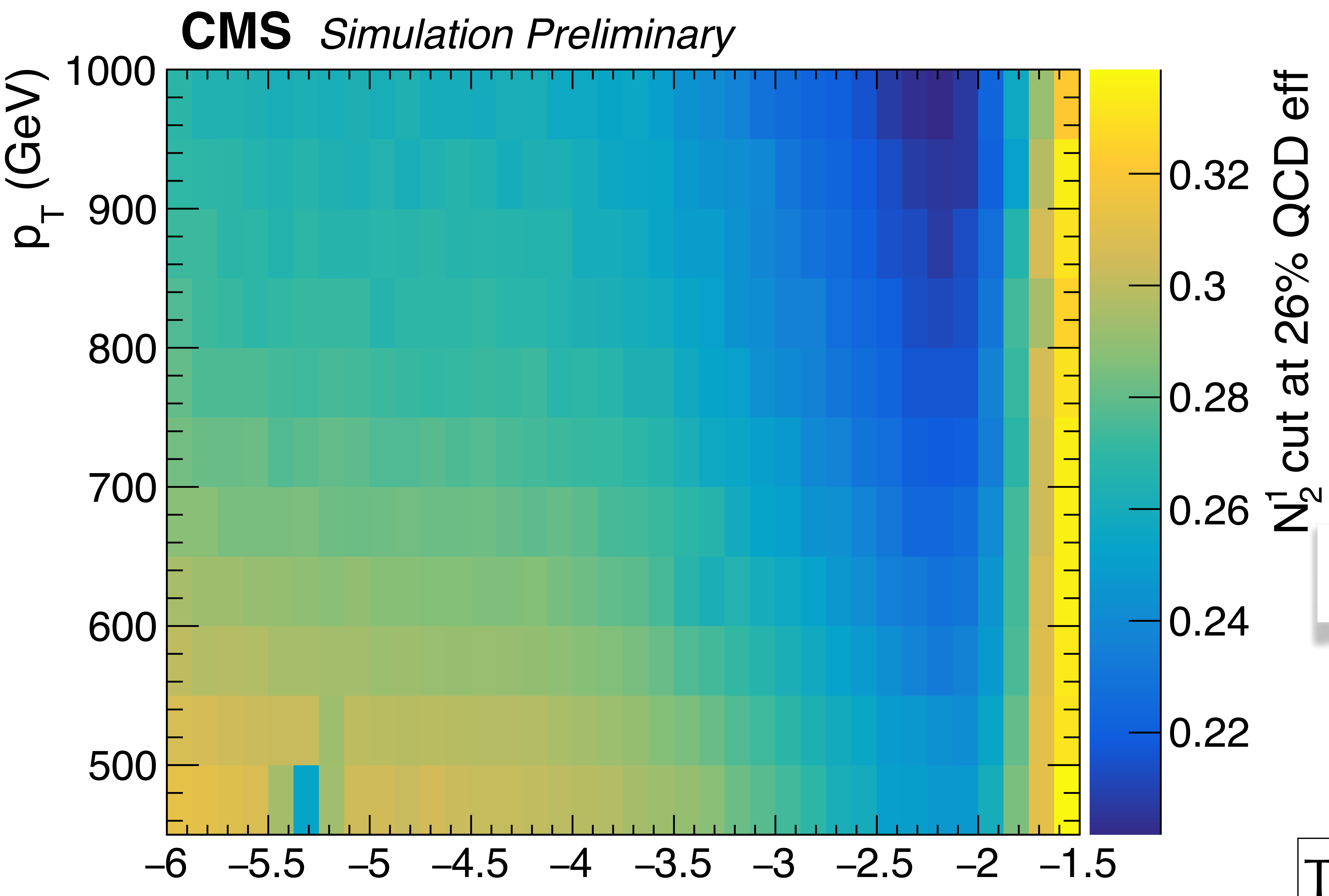
N^1_2 sculpts jet mass distribution

We use a **decorrelated** version for the background

$$N_2^{1,\mathrm{DDT}} = N_2^1 - N_2^1(\text{cut at } 26\% \text{ QCD eff.})$$

Jet Substructure

ArXiv:1603.00027
CMS-PAS-B2G-17-001
CMS-PAS-EXO-17-001



$\rho = \log(m_{SD}^2/p_T^2)$

$N^1{}_2$ sculpts jet mass distribution

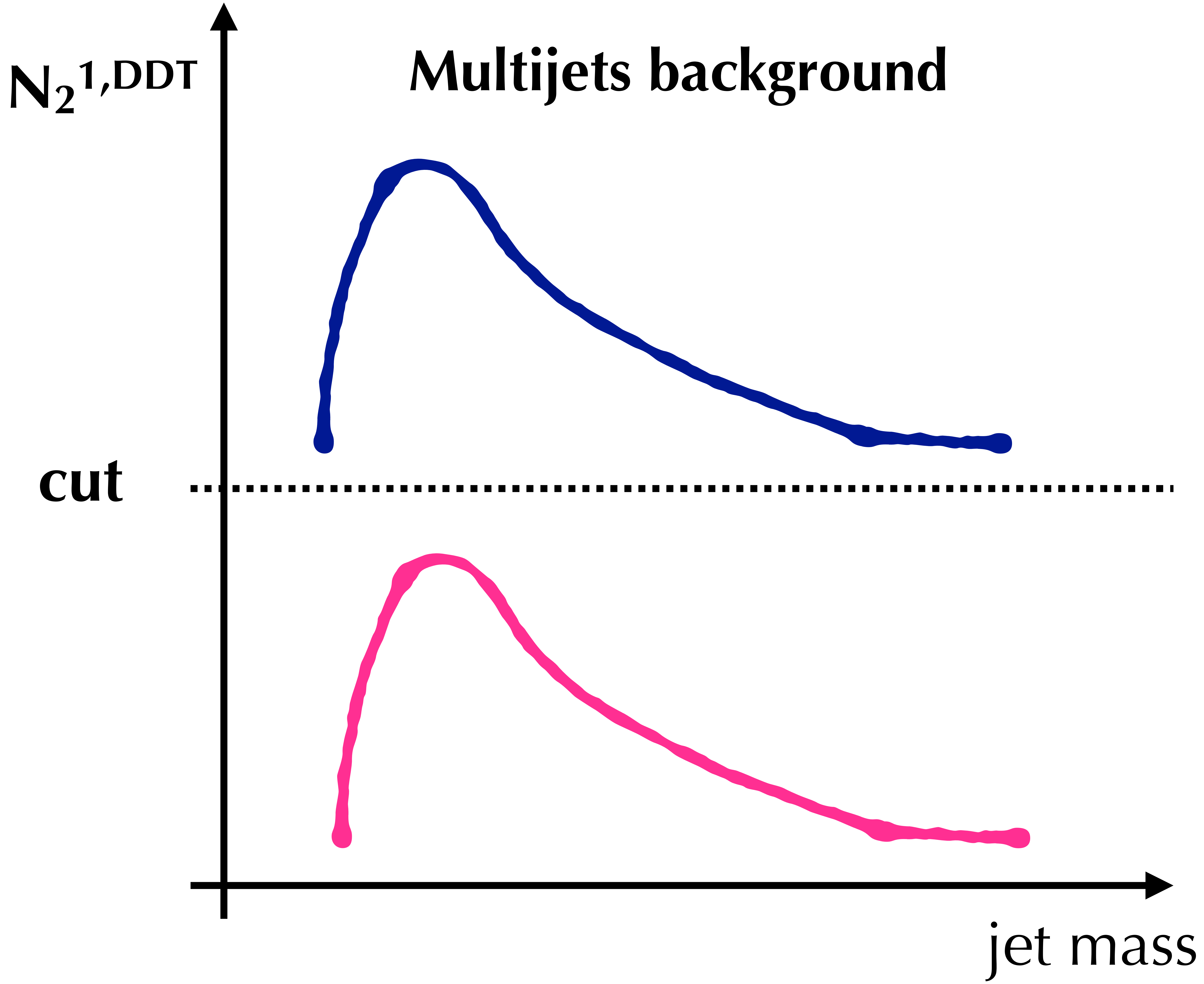
We use a **decorrelated** version for the background

$$N_2^{1,DDT} = N_2^1 - N_2^1(\text{cut at } 26\% \text{ QCD eff.})$$

The **scaling variable** for QCD jets ρ is used in the characterization of the correlation of jet substructure variable

Jet Substructure

ArXiv:1603.00027
CMS-PAS-B2G-17-001
CMS-PAS-EXO-17-001

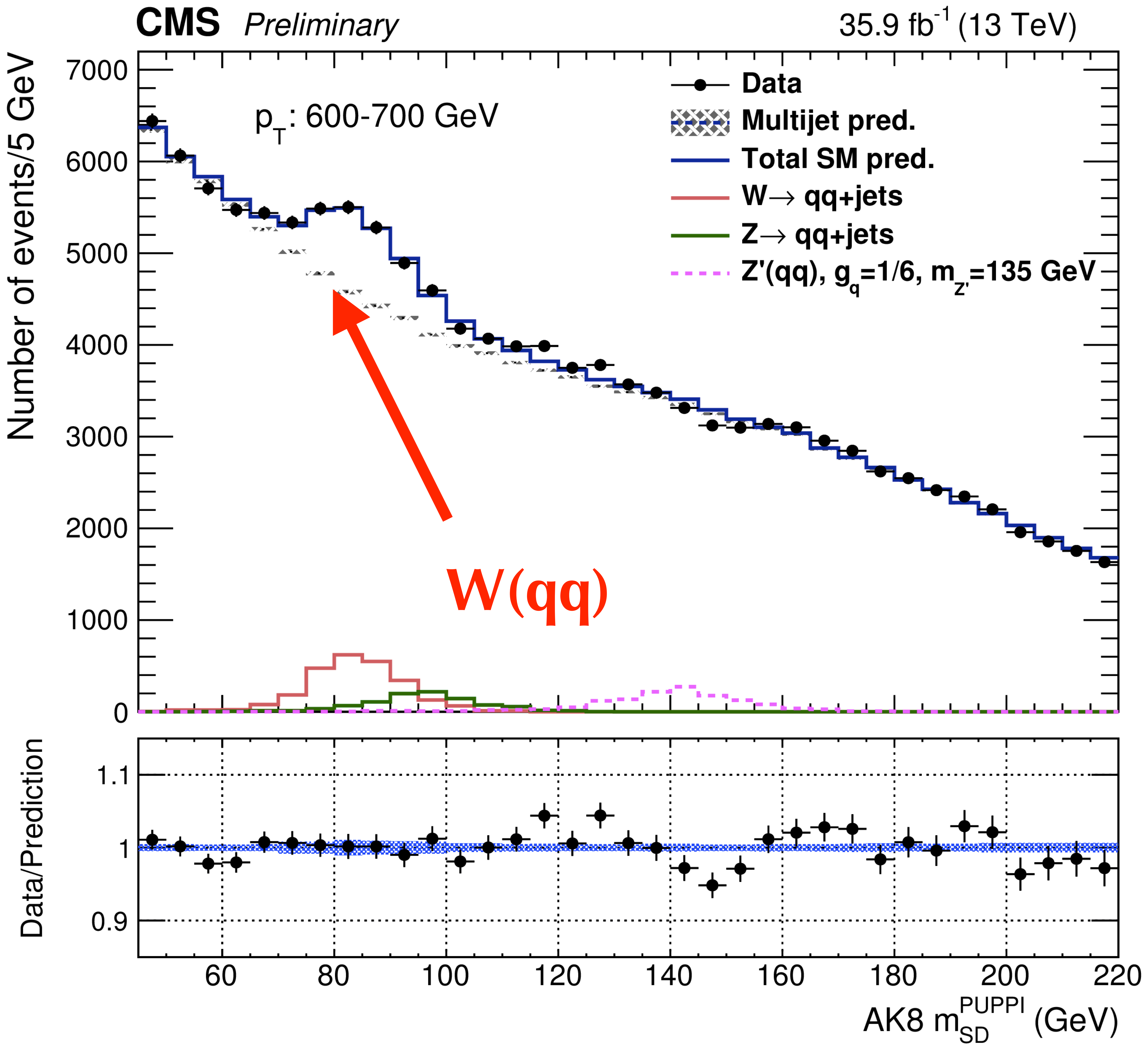


$N^1{}_2$ sculpts jet mass distribution

We use a **decorrelated** version for the background

Jet Substructure

ArXiv:1603.00027
CMS-PAS-B2G-17-001
CMS-PAS-EXO-17-001



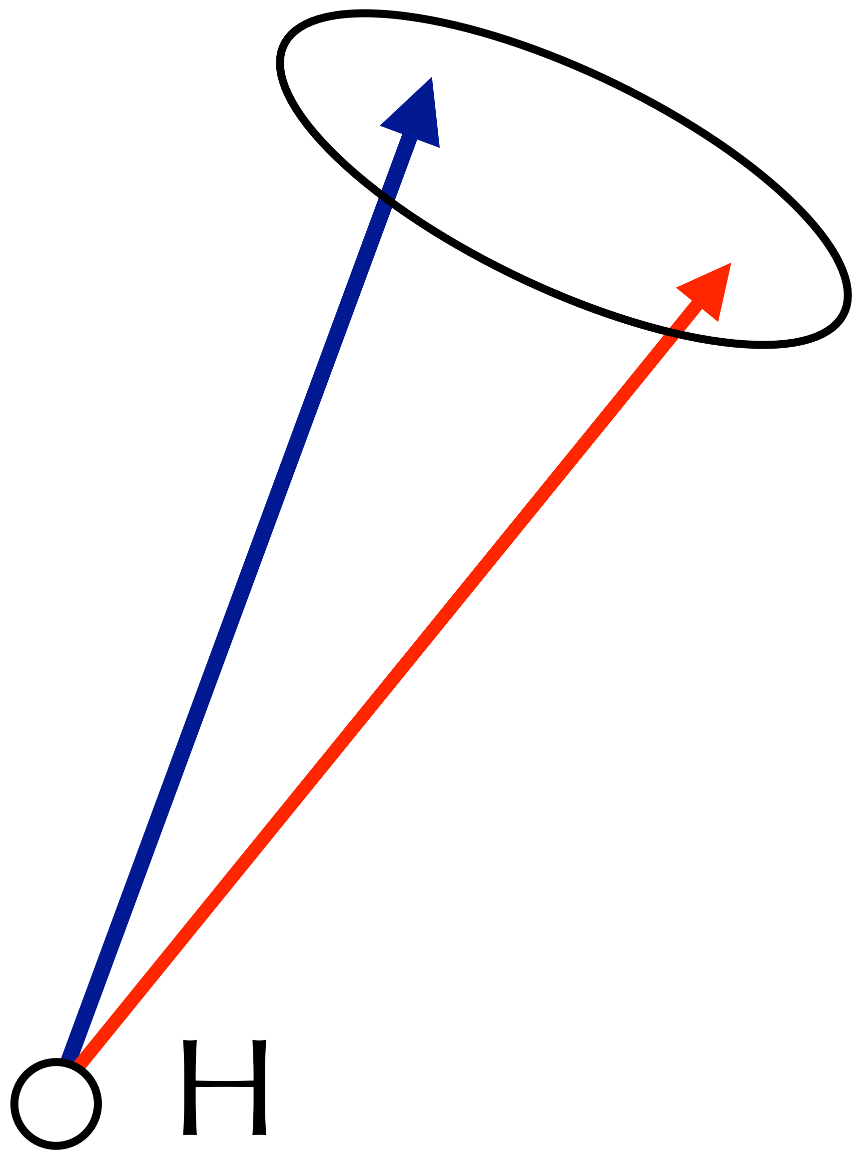
$N^1{}_2$ sculpts jet mass distribution

We use a **decorrelated** version for the background

Phys. Rev. Lett. 100 (2008) 242001

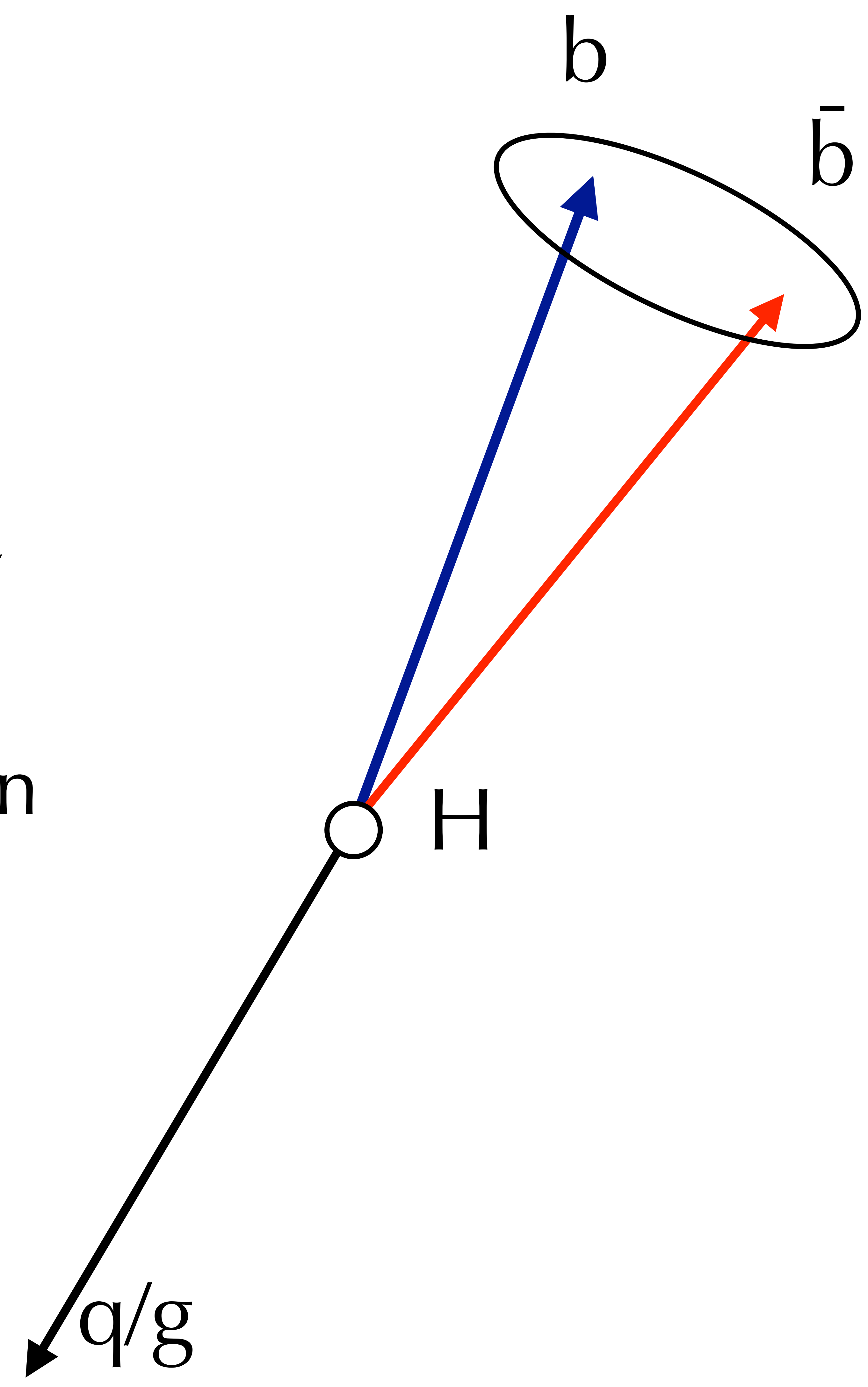
We conclude that subjet techniques have the potential to transform the high-p_T WH, ZH($H \to b\bar{b}$) channel into one of the best channels for discovery of a low mass Standard Model Higgs at the LHC.

now we search for <u>inclusive</u> H to bb with these new tools ...



Search for inclusive H to $b\bar{b}$

- We can access this process in the **boosted dijet** topology
- Use initial state jet to get above the trigger threshold
- Look for boosted **H boson in a single jet mass** distribution
  - Use the **Z boson as Standard Model** candle
    - b-tagging to disentangle W/Z

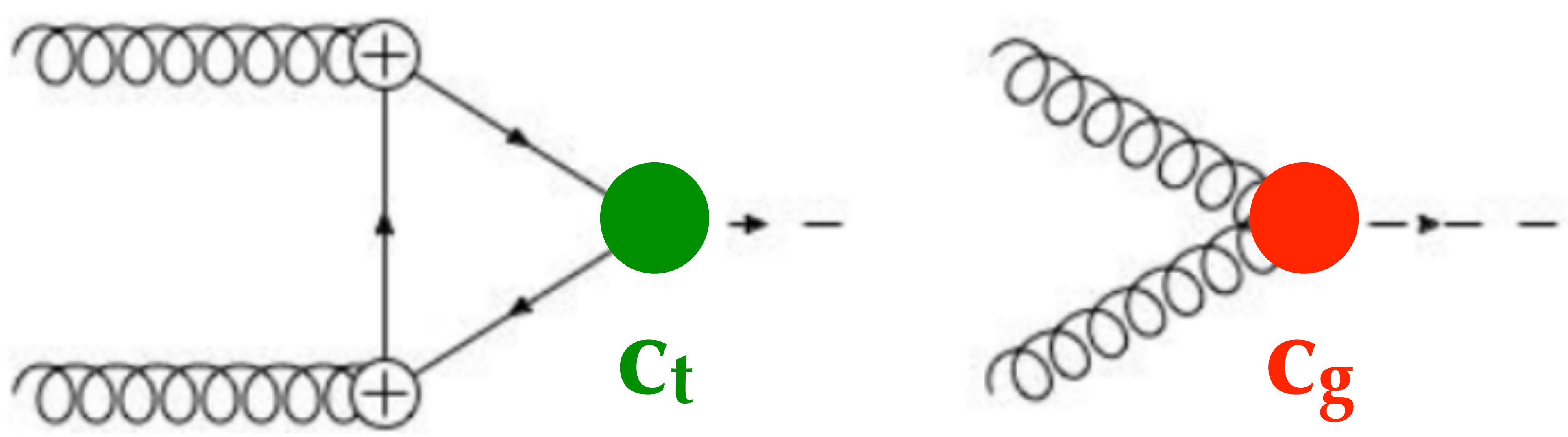


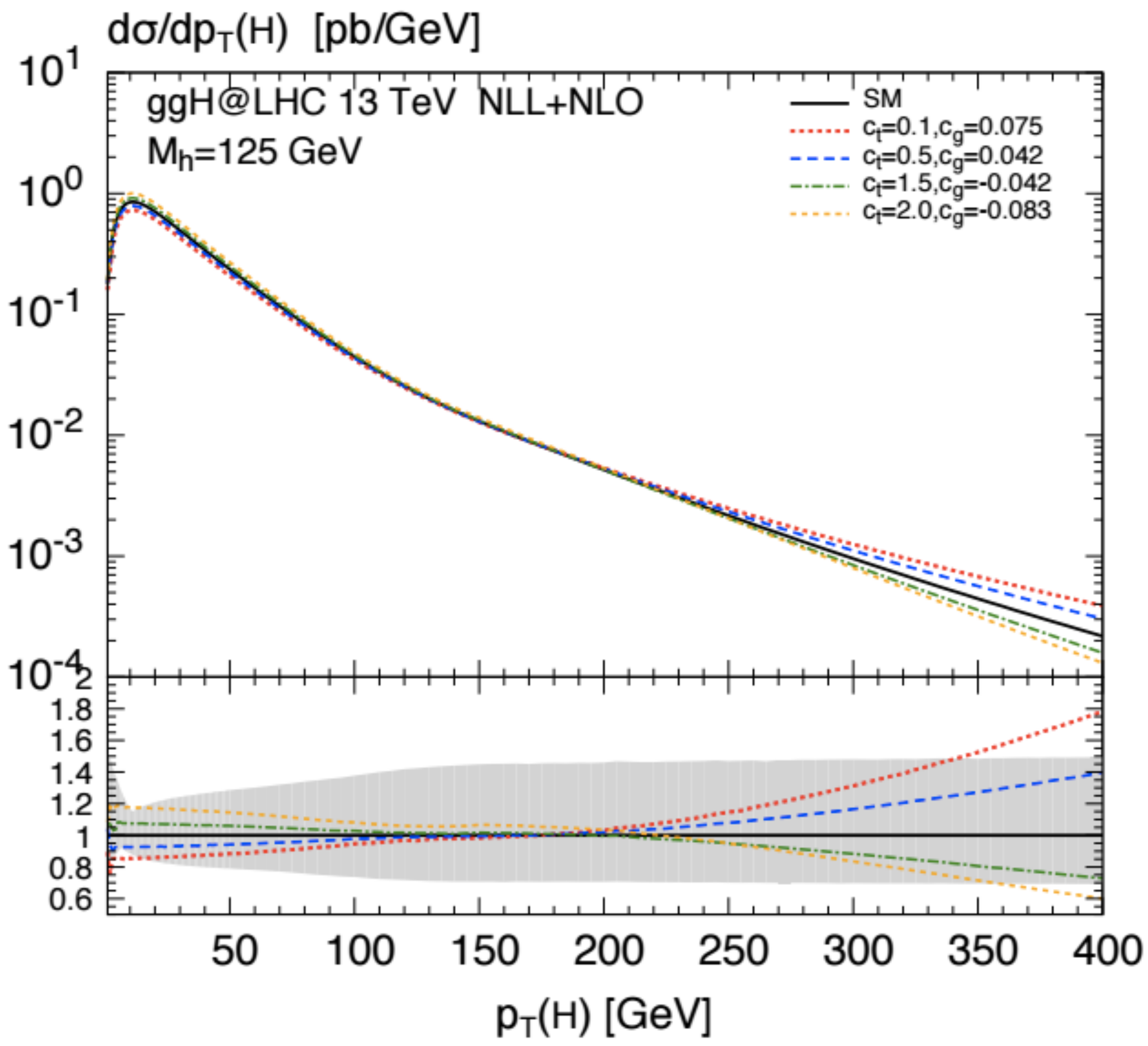
ArXiv:1612.00283
JHEP10(2016)123

Sensitivity to BSM

Probing Higgs couplings at high momentum transfer (Q) accesses large new physics energy scale (Λ)

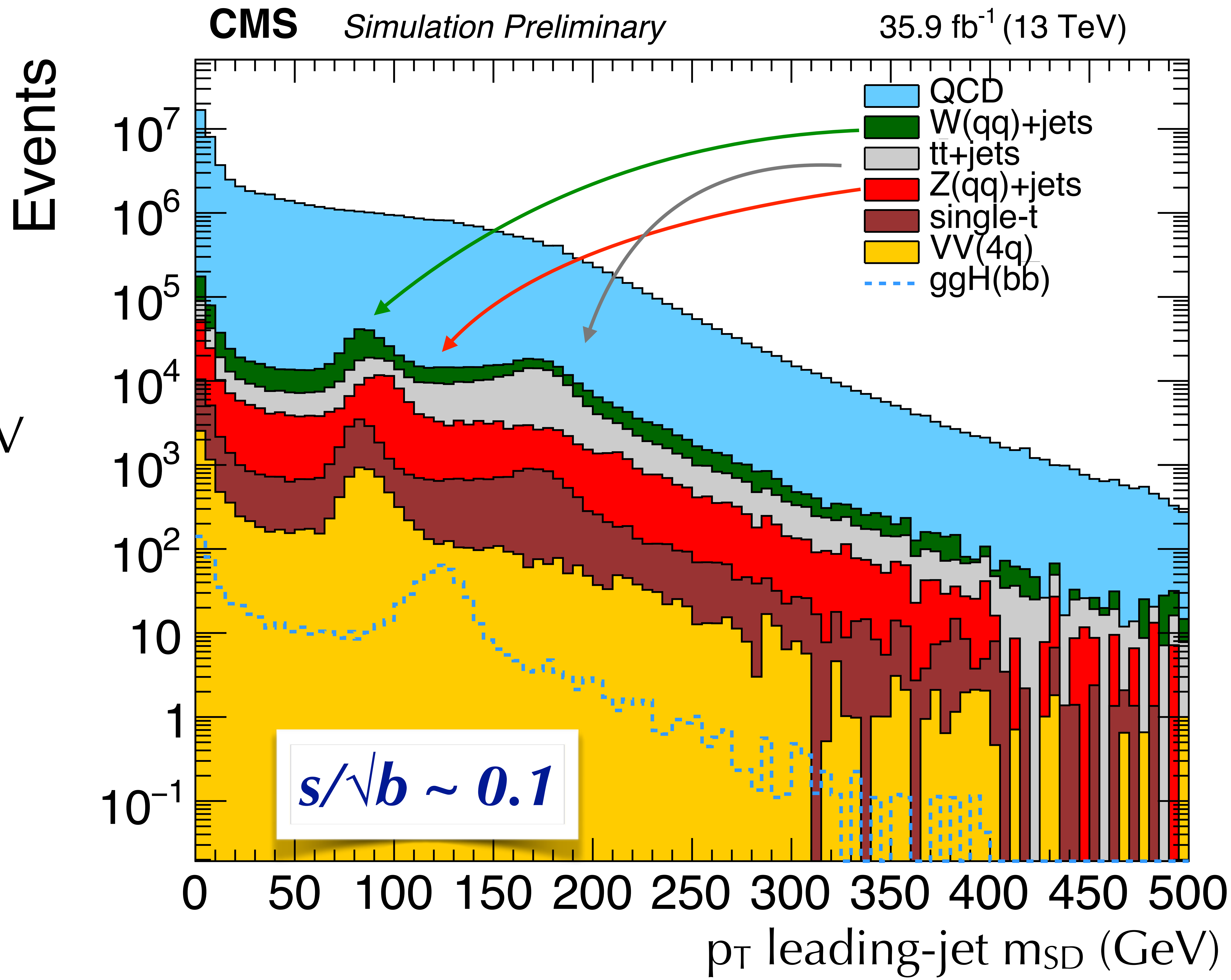
$$\Lambda_{\text{Sensitivity}} \approx \left(\frac{Q}{\text{vev}}\right)^2$$





Event Selection

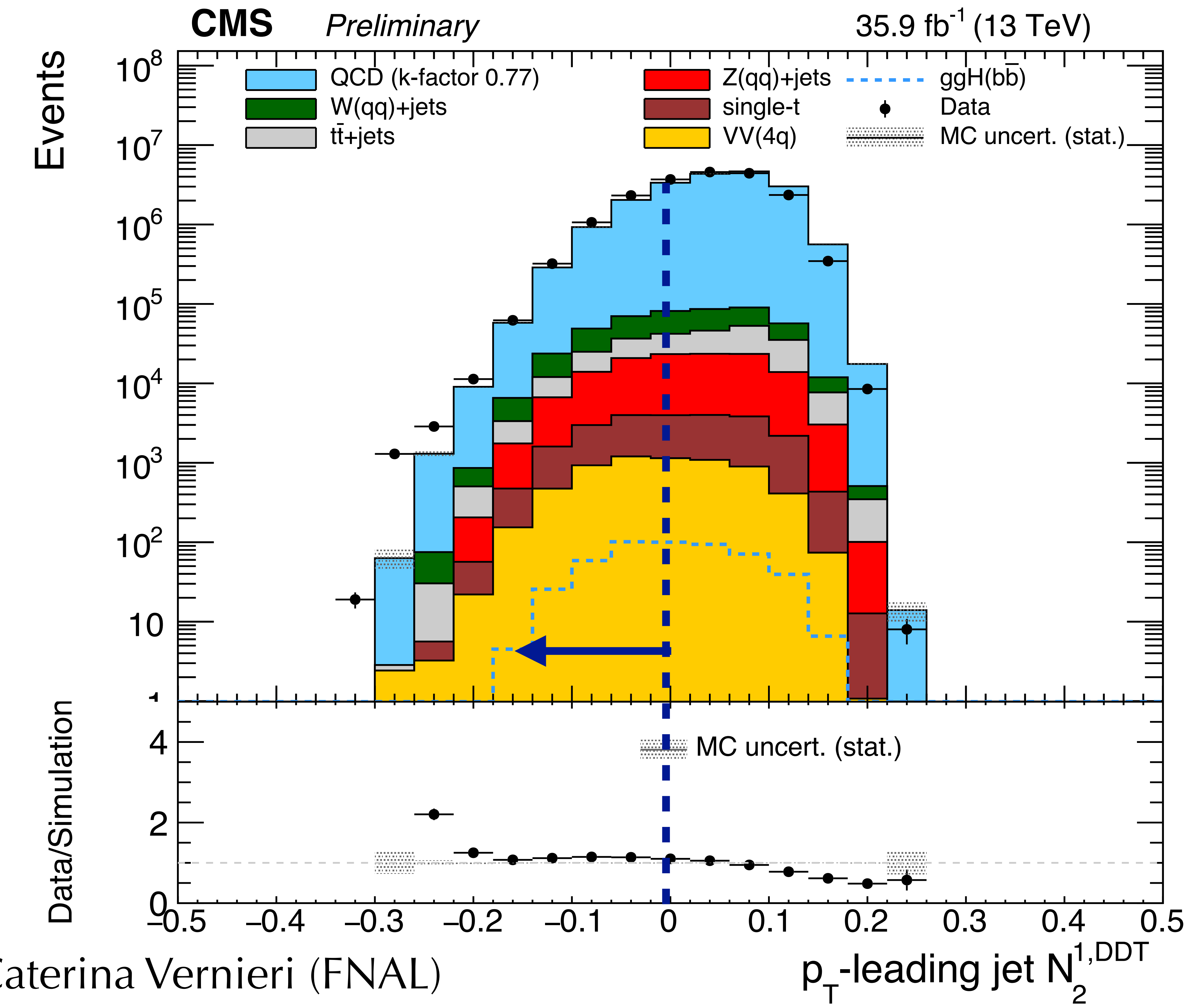
- **Online selection** asks for a high p_T single jet or large hadronic activities
  - $p_T > 360$ GeV or $\Sigma\, p_T > 800/900$ GeV
- Offline: **Highest p_T jet**:
  - $p_T > 450$ GeV $|\eta| < 2.5$,
  - jet soft drop mass (**m_{SD}**) > 40 GeV
- Lepton veto, E_T^{Miss} veto



Event Selection

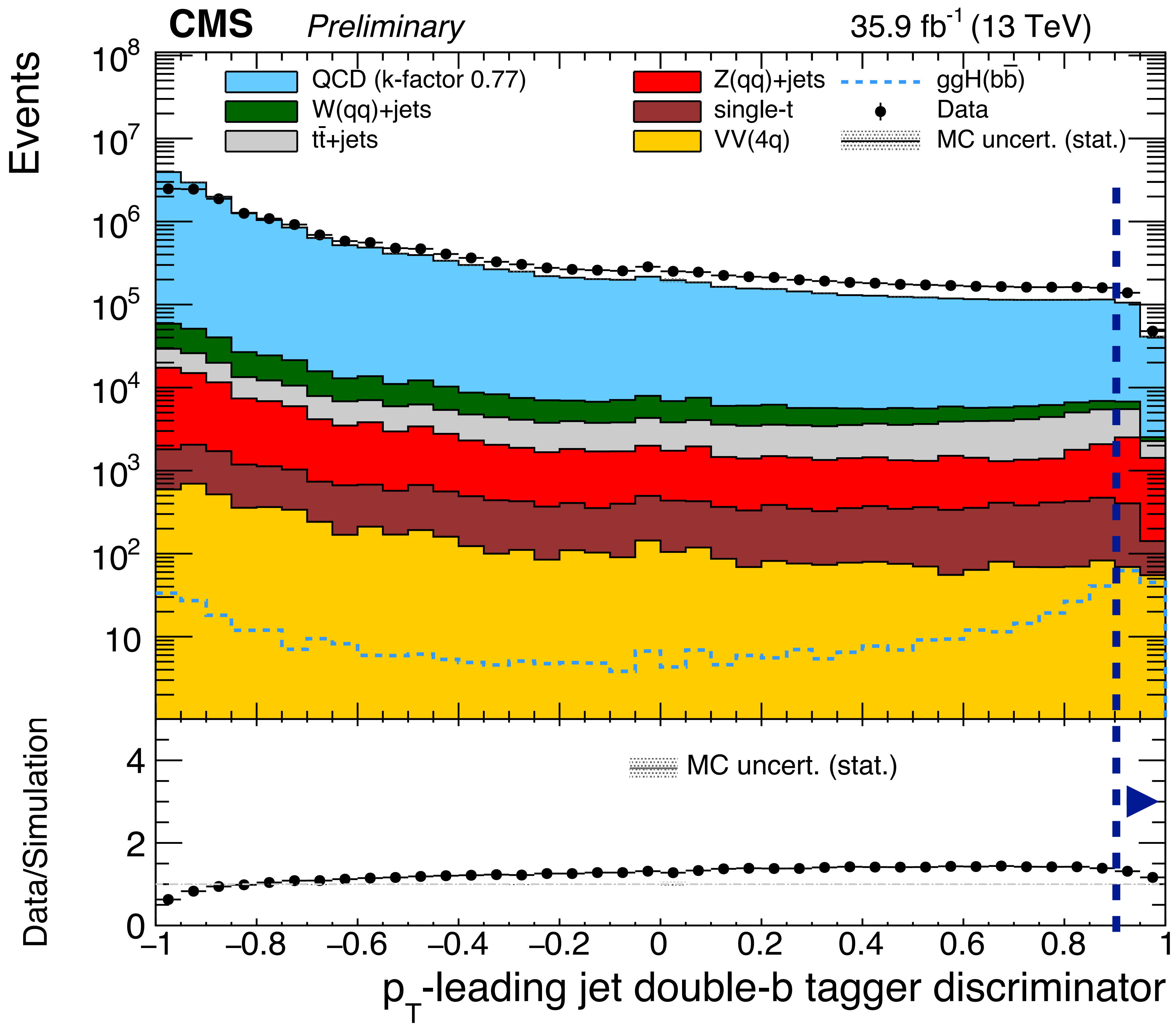
Substructure: two prongs discrimination

~50% sig. efficiency; 26% background efficiency



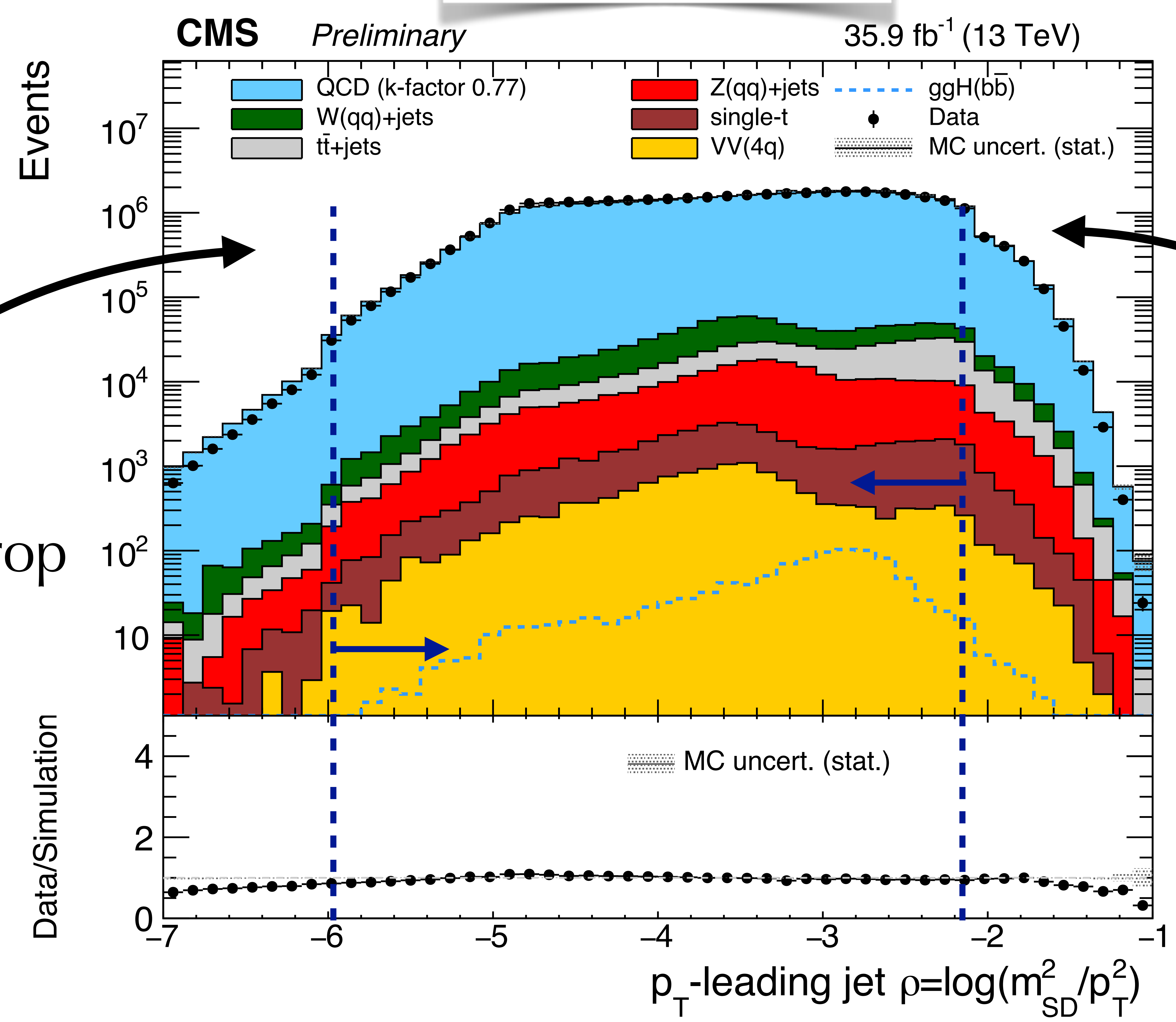
double-b tagger

~30% sig. efficiency; 1% background efficiency (tight working point)



Event Selection

$-6 < \rho < -2.1$

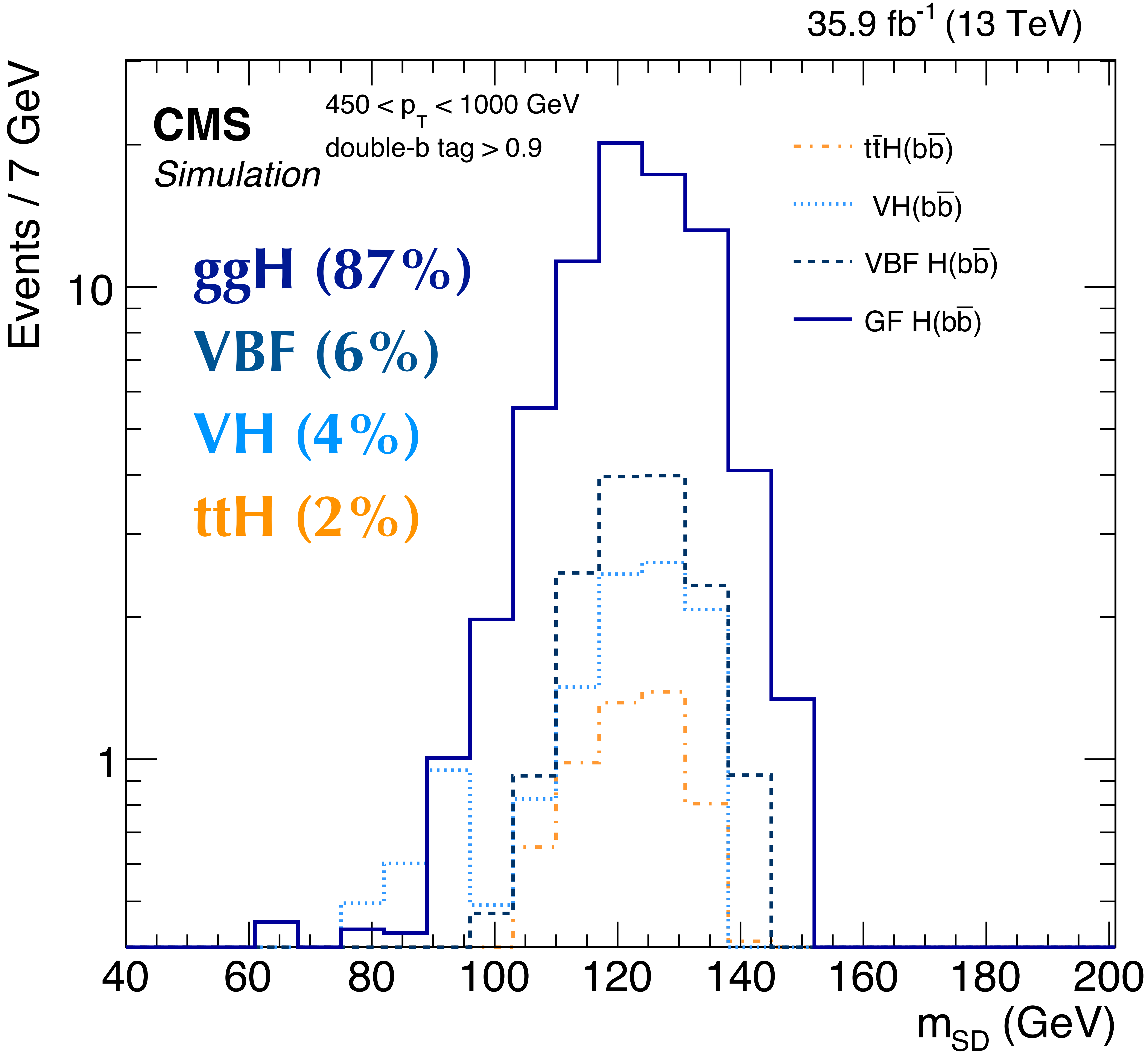


non-perturbative regime of the soft drop mass calculation

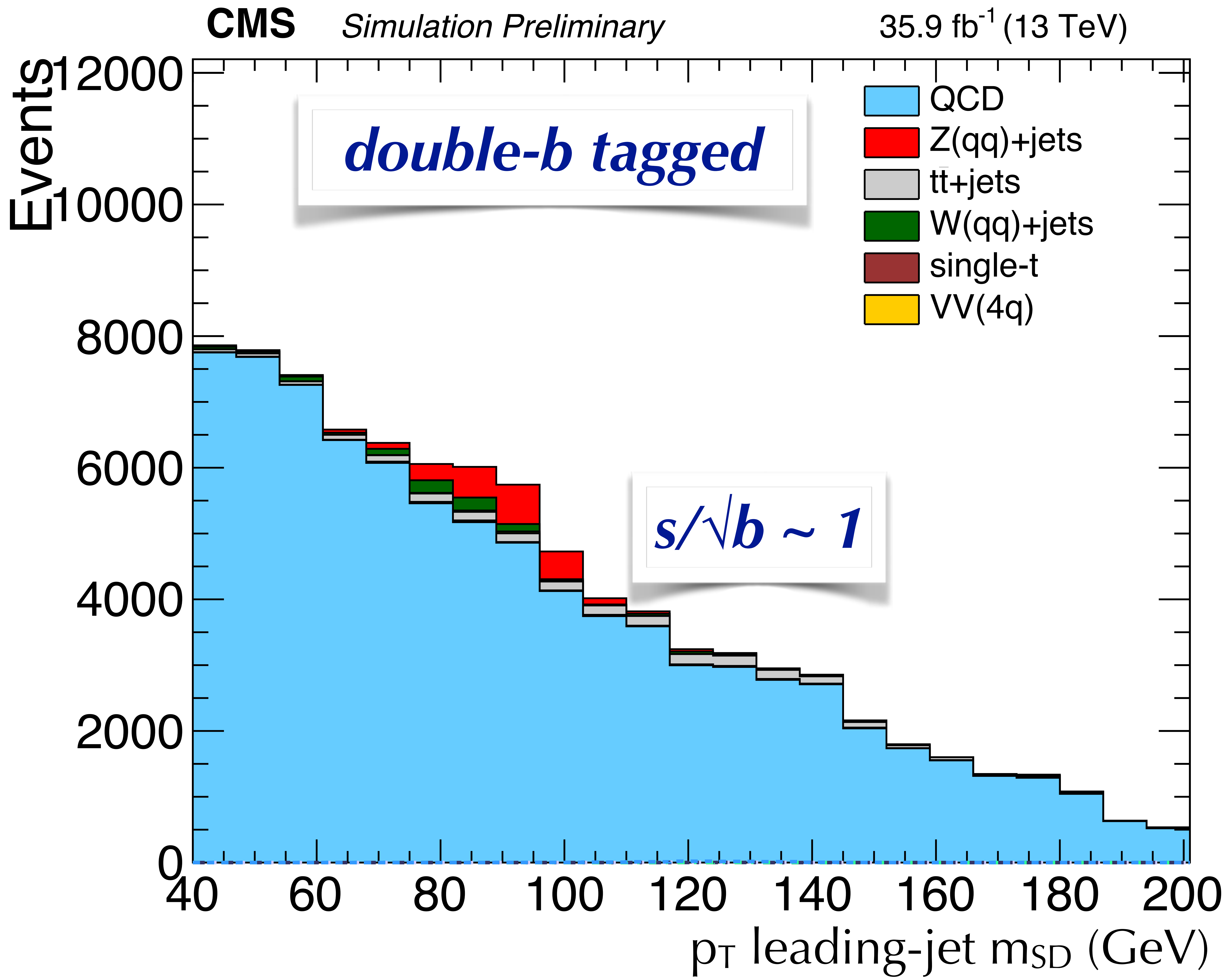
finite cone effects from the AK8 jet clustering

After all selections

- Analysis is inclusive in Higgs production mode
- Dominant contribution in signal region is ggH



Background composition



Backgrounds

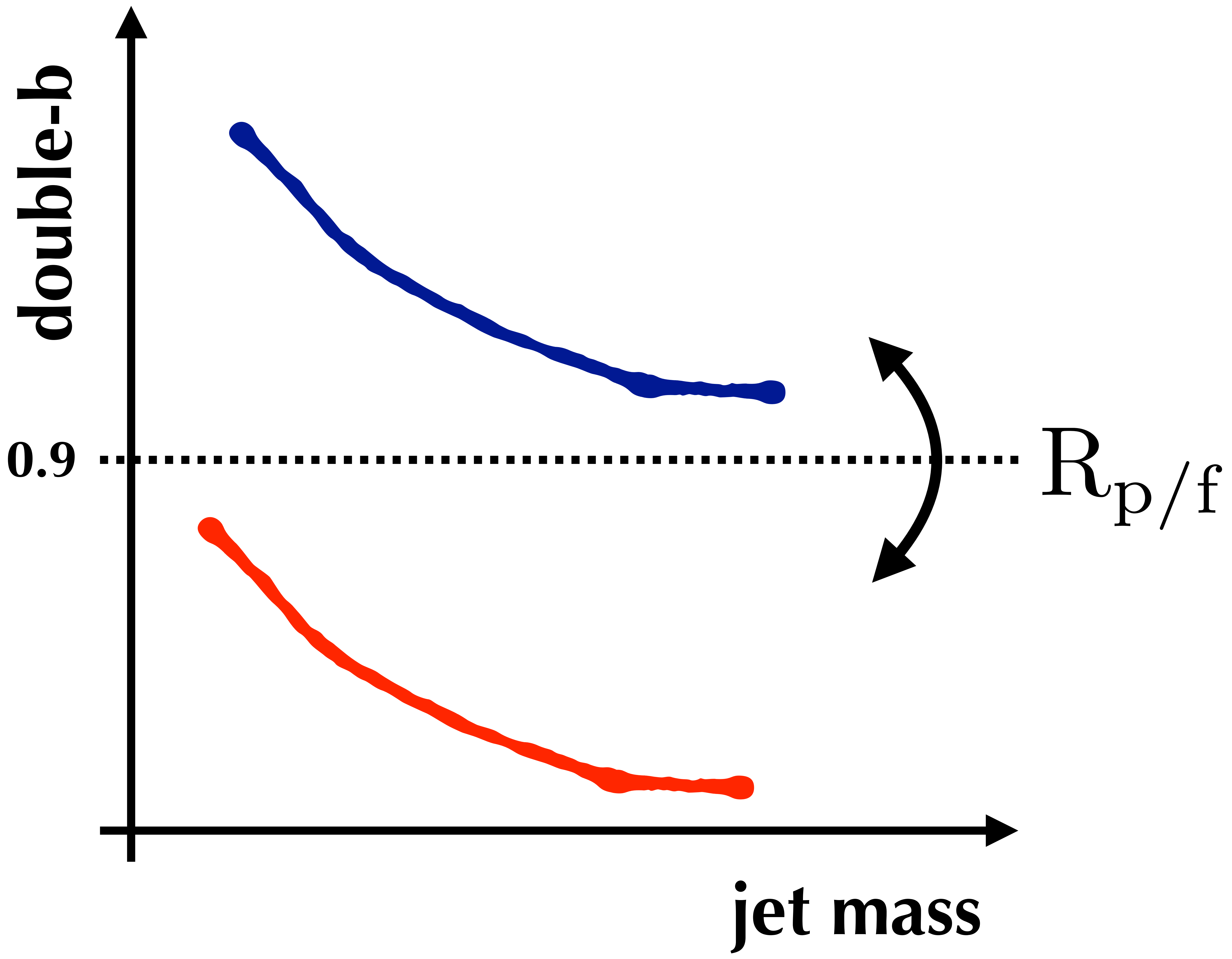
- QCD (~**90%**)
- tt+jets (**3%**), normalization from a dedicated control region

} **data**

- W/Z+jets (**5%**)
- single-t, VV (<**1%**)

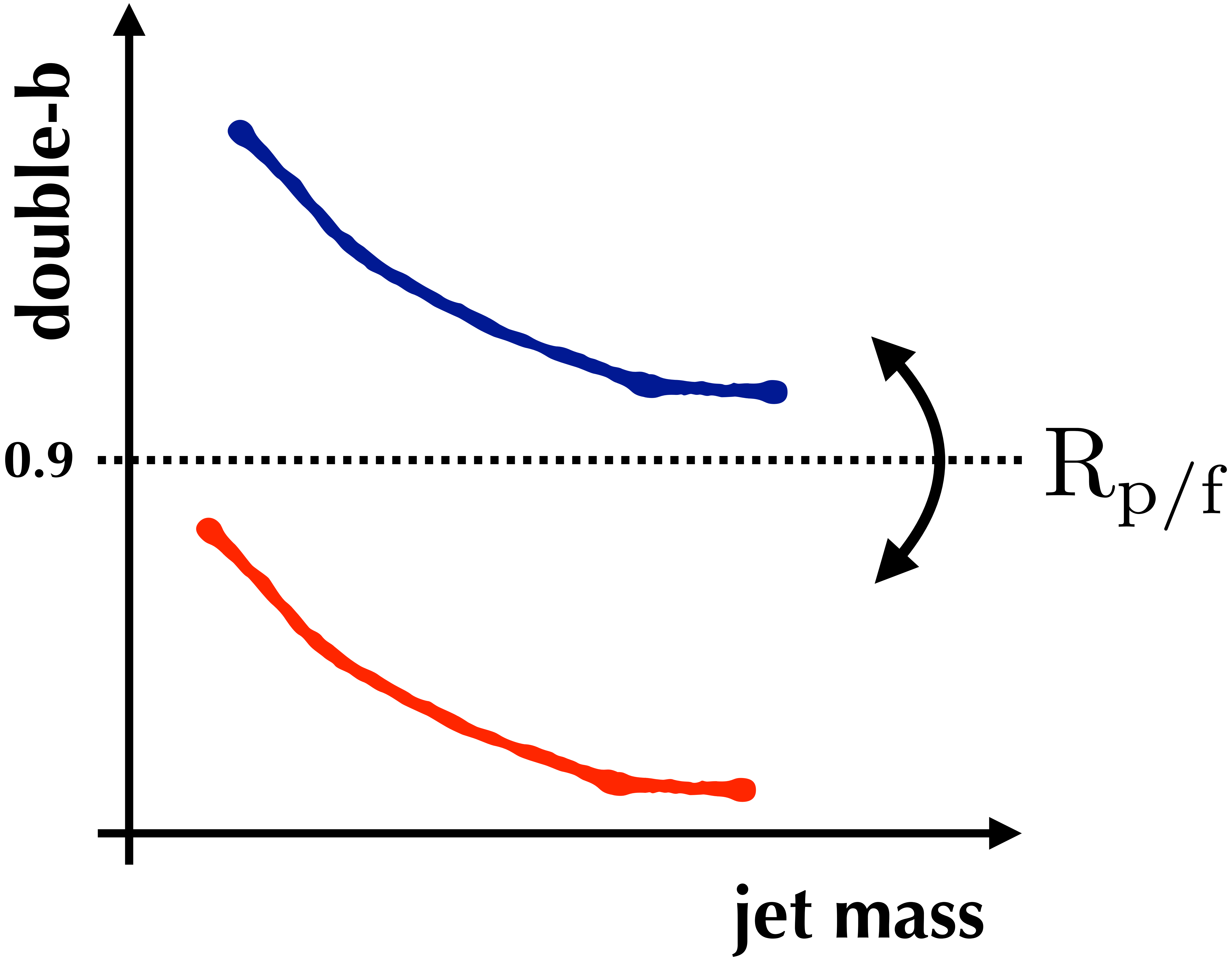
} **MC**

Multijets background prediction

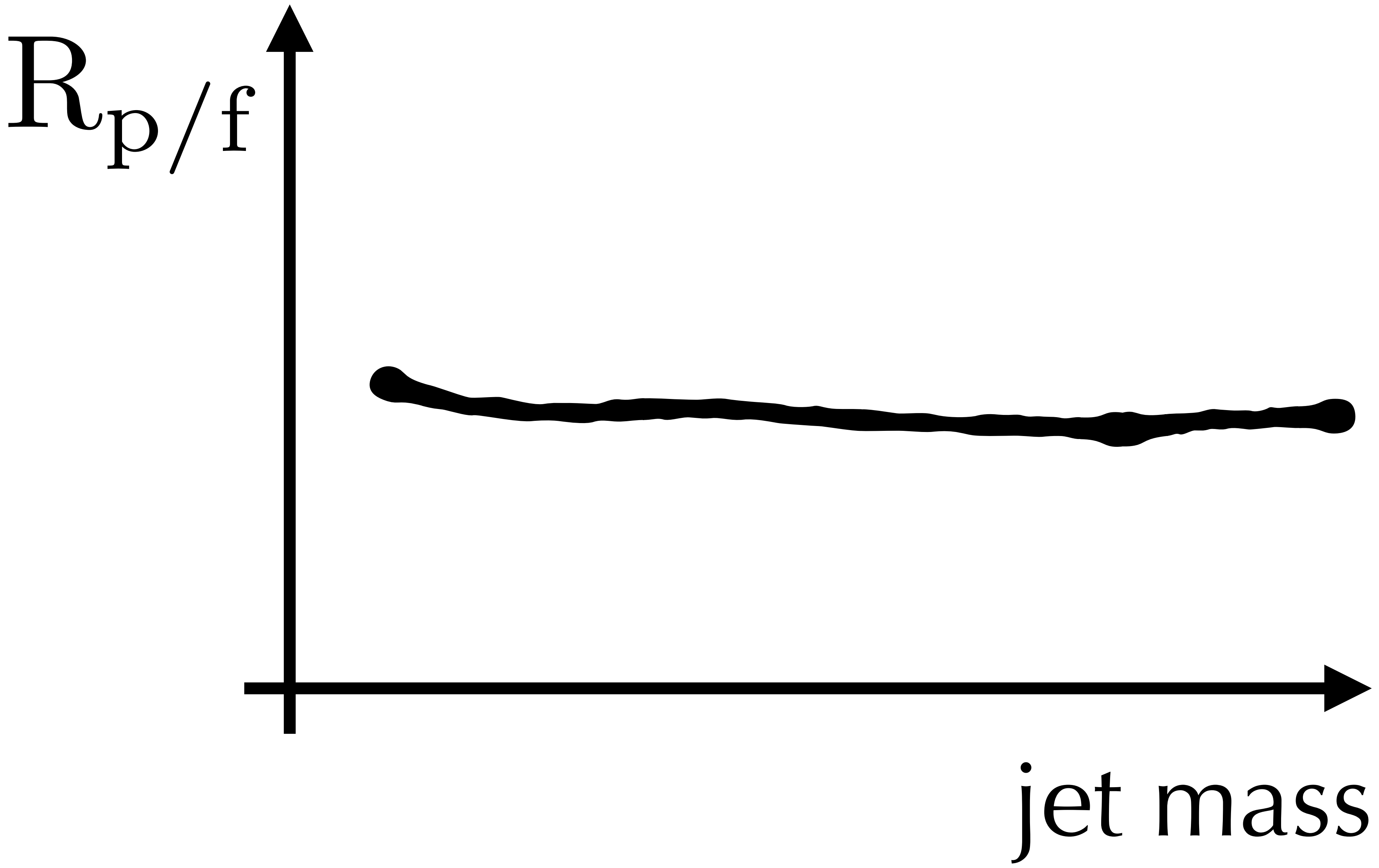


- Jet mass shape for multijets events is derived in data
- From events "failing" the b-tag requirement with a ***transfer factor***

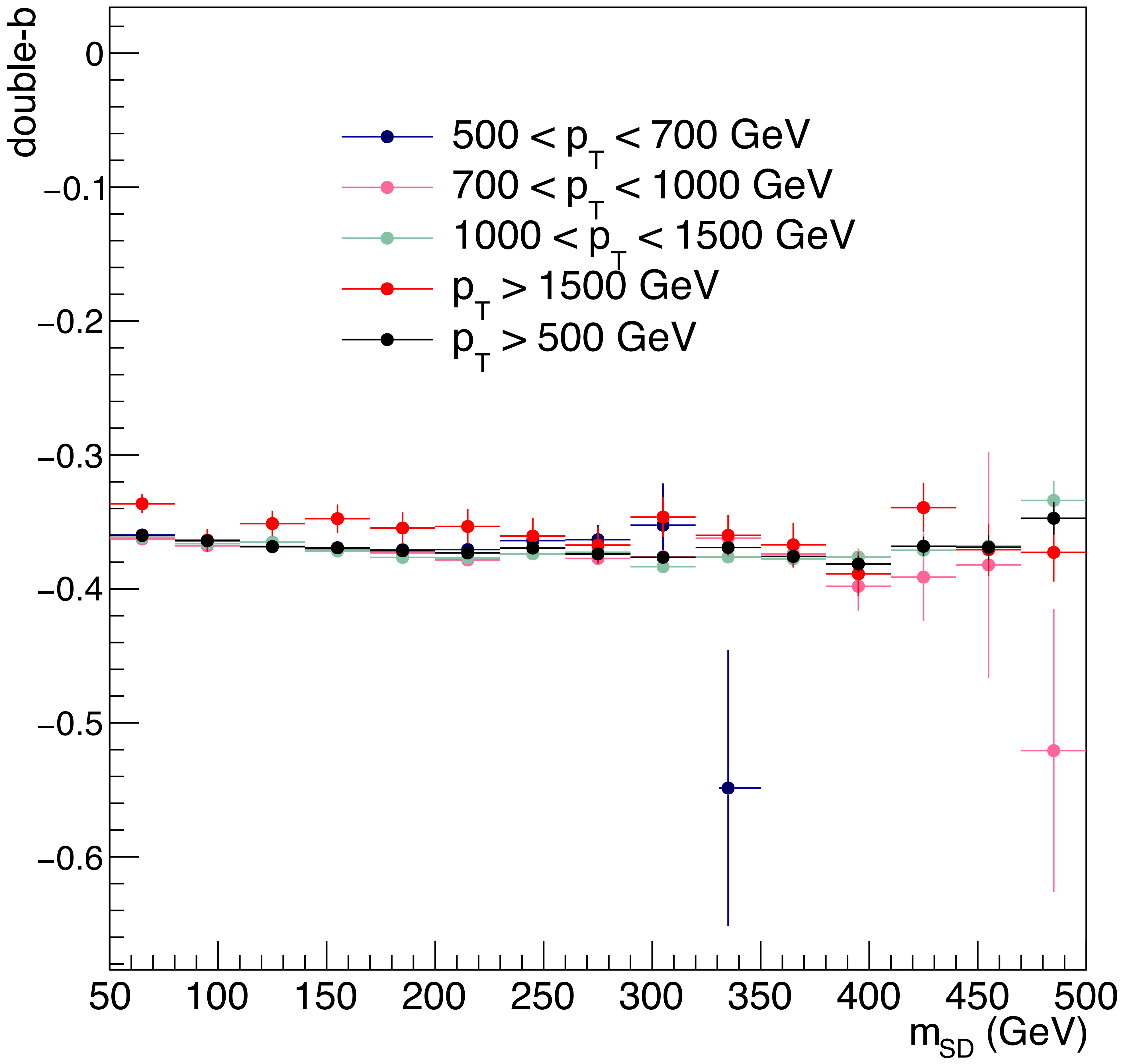
Multijets background prediction



- Jet mass shape for multijets events is derived in data
- From events "failing" the b-tag requirement with a ***transfer factor***



Multijets background prediction



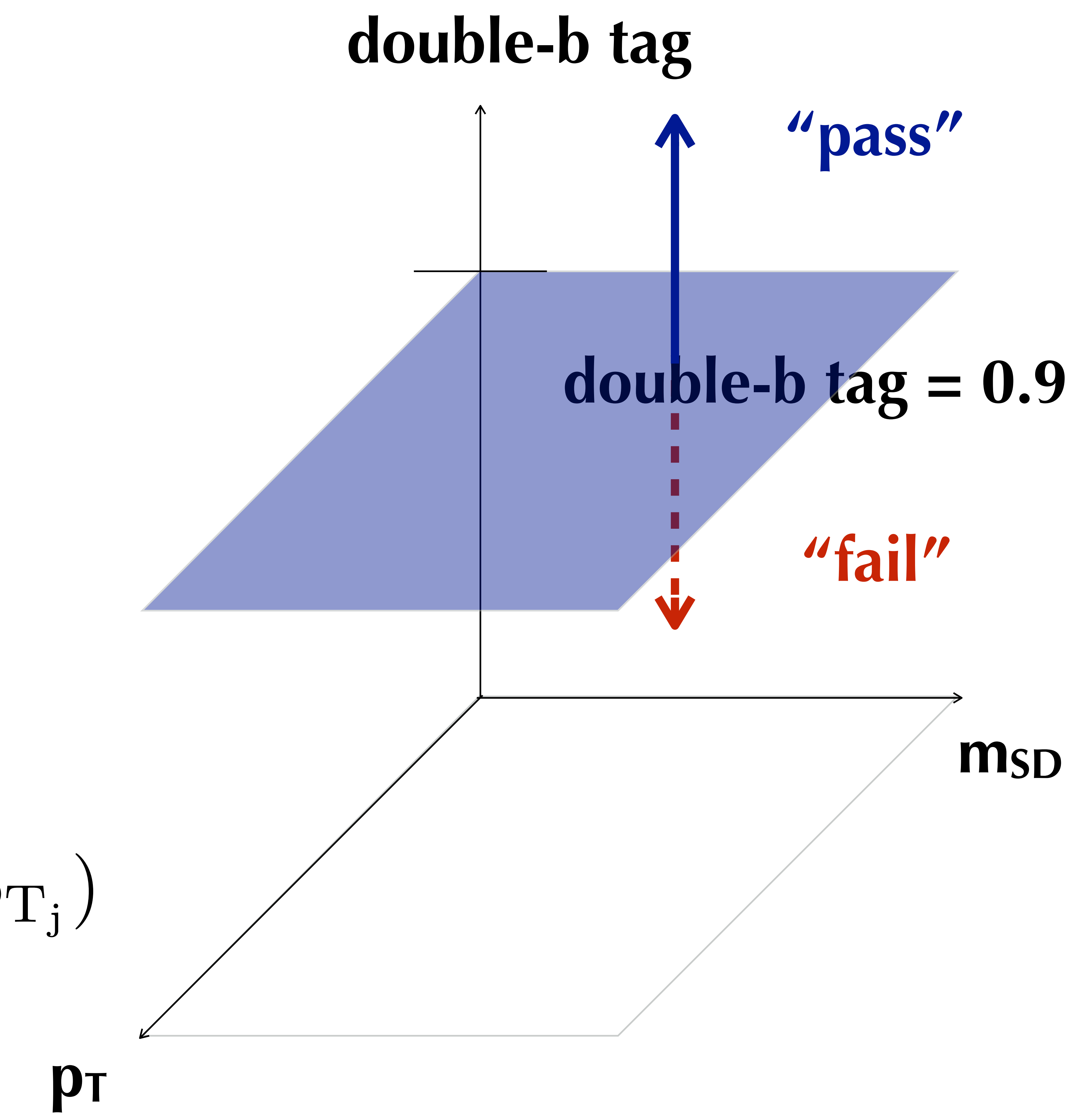
- Jet mass shape for multijets events is derived in data
- From events "failing" the b-tag requirement with a ***transfer factor***
  - as function of the **jet mass and p_T**

Transfer factor

- If the double-b tagger were completely uncorrelated from jet p_T and m_{SD}, the transfer factor would be flat
  - **Taylor expand** $R_{p/f}$ in ρ and p_T

$$N^{QCD}_{pass}(m_{SD}, p_T) = R_{p/f}(\rho, p_T) \times N^{QCD}_{fail}(m_{SD}, p_T)$$

$$N^{QCD}_{pass}(m_{SD_i}, p_{T_j}) = \left(\sum_{k,\ell} a_{kl} \rho^{kl}_{ij} p^{\ell}_{T_j} \right) \times N^{QCD}_{fail}(m_{SD_i}, p_{T_j})$$



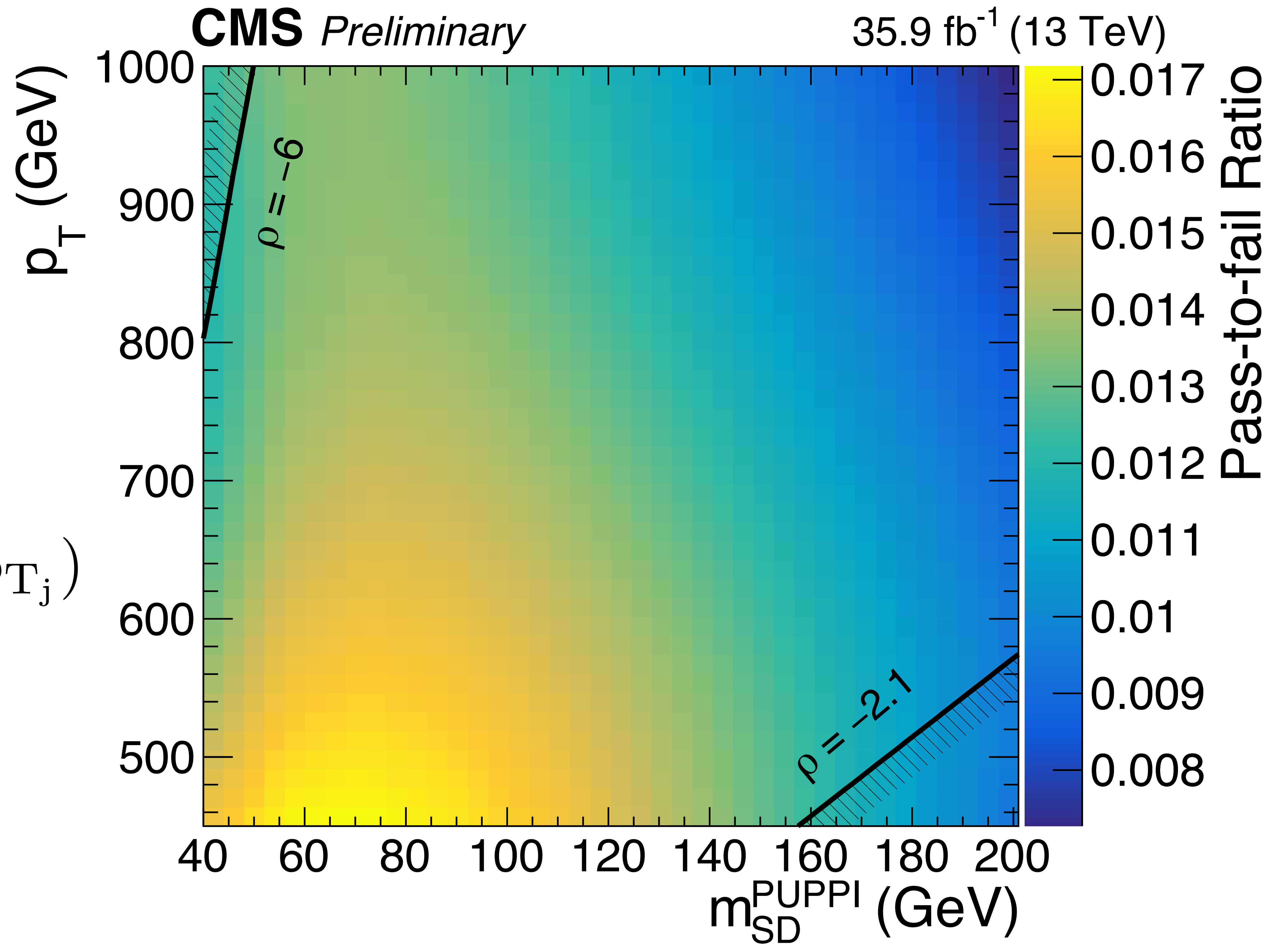
- **23 bins** in $\mathbf{m_{SD}}$ from **40 to 201 GeV** and **6 bins** in $\mathbf{p_T}$ from **450 to 1000 GeV**

Transfer factor determination

- Signal extraction and background estimation performed **simultaneously**

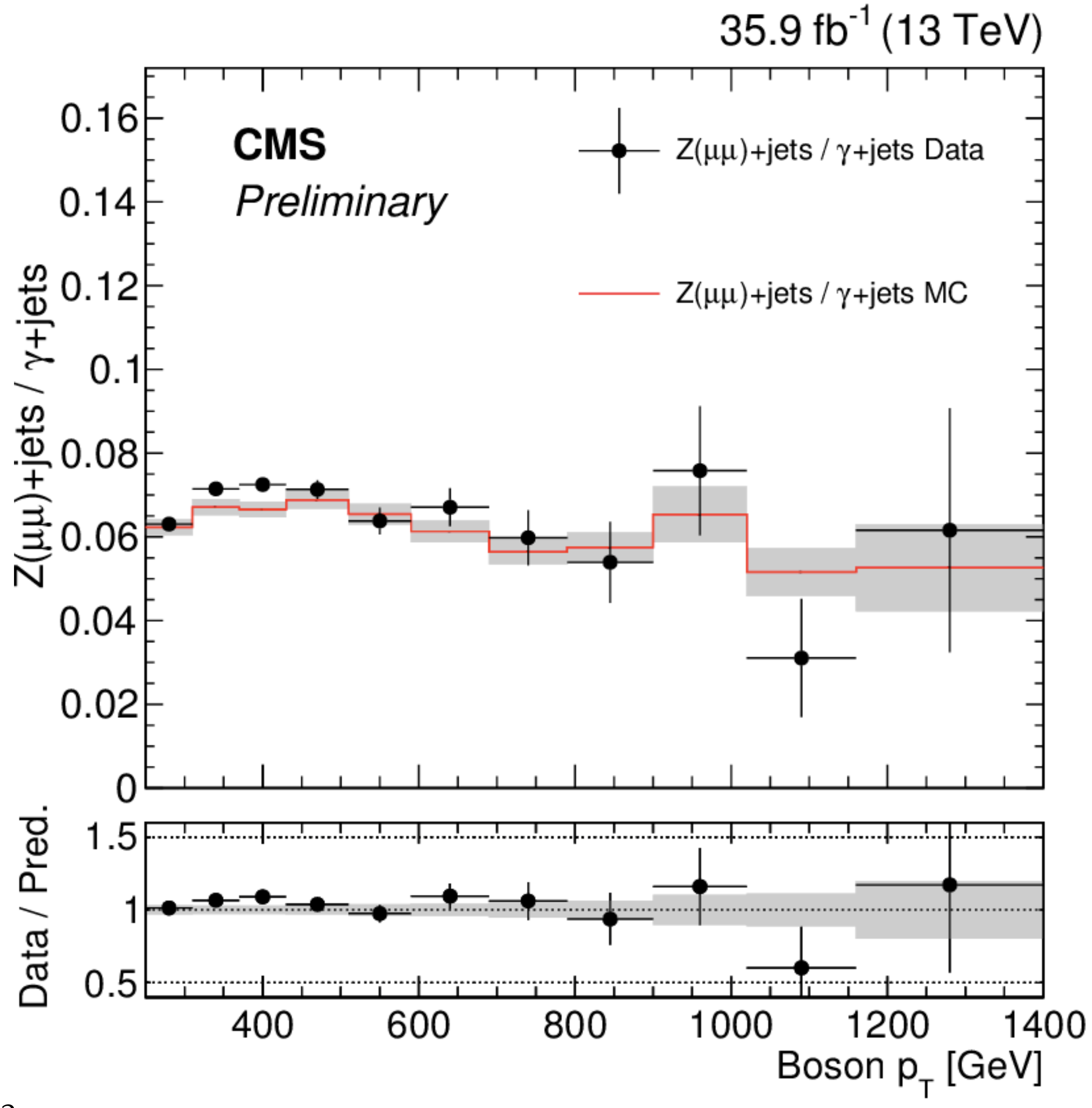
$$\mathcal{L}(\text{data}|\mu,\boldsymbol{\theta}) = \prod_{i,j} \text{Poisson}\left(N^{\text{data}}_{\text{fail},i,j} | N^{\text{QCD}}_{\text{fail},i,j} + N^{t\bar{t}}_{\text{fail},i,j} + N^{V}_{\text{fail},i,j} + \mu\ N^{H(b\bar{b})}_{\text{fail},i,j}\right)$$
$$\times \prod_{i,j} \text{Poisson}\left(N^{\text{data}}_{\text{pass},i,j} | N^{\text{QCD}}_{\text{pass},i,j} + N^{t\bar{t}}_{\text{pass},i,j} + N^{V}_{\text{pass},i,j} + \mu\ N^{H(b\bar{b})}_{\text{pass},i,j}\right)$$

$$\mathrm{N^{QCD}_{pass}(m_{SD_i}, p_{T_j})} = \left(\sum_{\mathrm{k},\ell} \mathrm{a_{kl}} \rho^{\mathrm{kl}}_{\mathrm{ij}} \mathrm{p}^{\ell}_{\mathrm{T_j}}\right) \times \mathrm{N^{QCD}_{fail}(m_{SD_i}, p_{T_j})}$$



- The coefficients a_{kl} are determined from the fit
- Based on F-test, a second order polynomial in ρ and first order in p_T is used.

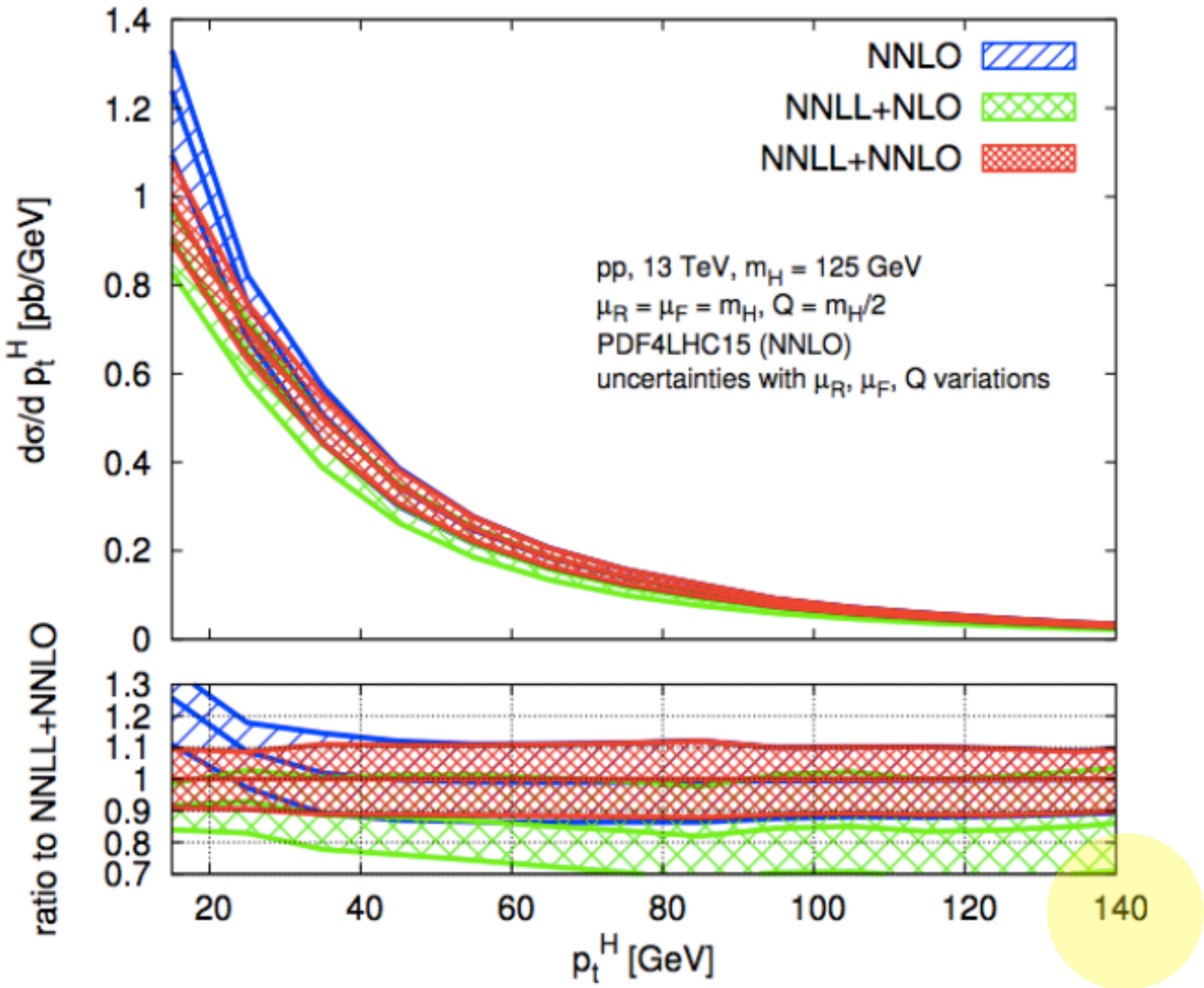
W/Z+jets simulation



- LO simulations for the W/Z+jets are corrected using **p_T-dependent** :
  - **NLO QCD** k-factors
  - **NLO electroweak** k-factors
- Associated uncertainties are 10% (QCD) and 15-35% (EWK)
- Additional systematic of 5% accounts for potential differences between the W and Z higher order corrections.

ggH simulation

- Other CMS Higgs results use **Powheg** : 1 jet + $m_t = \infty$
- We want to account for both the effects of higher order corrections and for the finite top mass
  - No real NLO + finite top mass calculation available in the literature above $p_T^H > 300$ GeV



<< 1000 GeV

ggH simulation

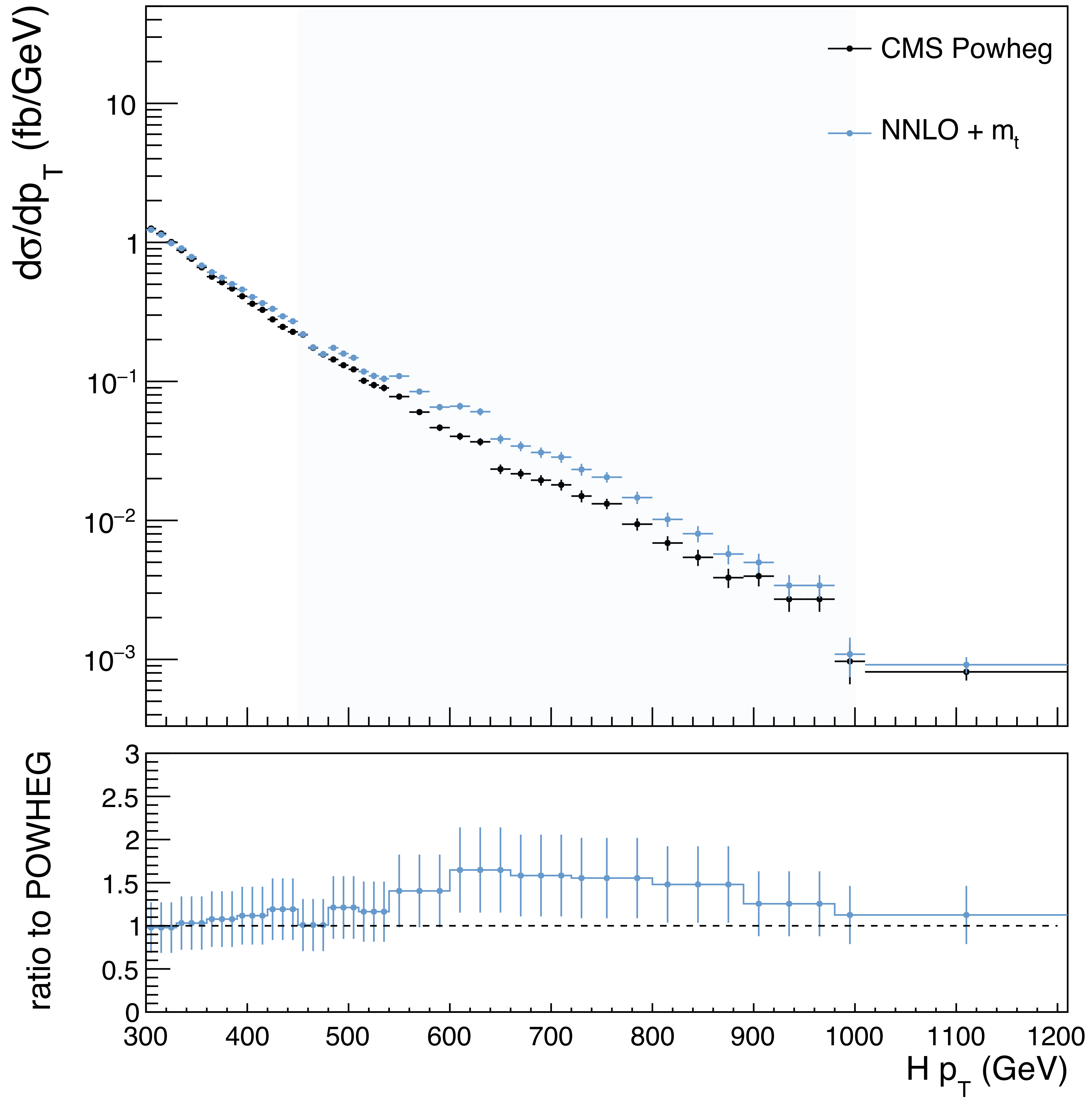
ArXiv:1410.5806, ArXiv:1609.00367
ArXiv:1408.5325, ArXiv:1302.6216
ArXiv:1504.07922, ArXiv:1505.03893
ArXiv:1610.07922

- Other CMS Higgs results use **Powheg** : 1 jet + $m_t = \infty$
- We want to account for both the effects of higher order corrections and for the finite top mass
  - No real NLO + finite top mass calculation available in the literature above $p_T^H > 300$ GeV
- A multi-correction approach is adopted
  - LO H+0-2jet, finite m_T
  - NLO H+1jet finite m_t up to $1/m_t^4$ expansion
  - NNLO H+1jet, $m_T = \infty$

$$\mathrm{GF\ H(NNLO + m_t) = (1\ jet\ m_t = \infty) \times \frac{MG\ LO\ 0-2\ jet\ m_t}{(1\ jet\ m_t = \infty)} \times \frac{NLO\ 1\ jet\ m_t}{LO\ 1\ jet\ m_t} \times \frac{NNLO\ 1\ jet\ m_t = \infty}{NLO\ 1\ jet\ m_t = \infty}}$$

ggH p_T reweighting finite m_t + NNLO

- Estimate k-factor of ~1.3 for Higgs $p_T > 450$ GeV
- Two factorized systematic uncertainties:
  - 30% normalization
  - 30% change in slope (no effect on overall norm.)
- This is the **first time an (approximate) NLO H+0,1,2 jet merged with finite top mass** is attempted



Results are stable under these variations and also provided without p_T reweighting

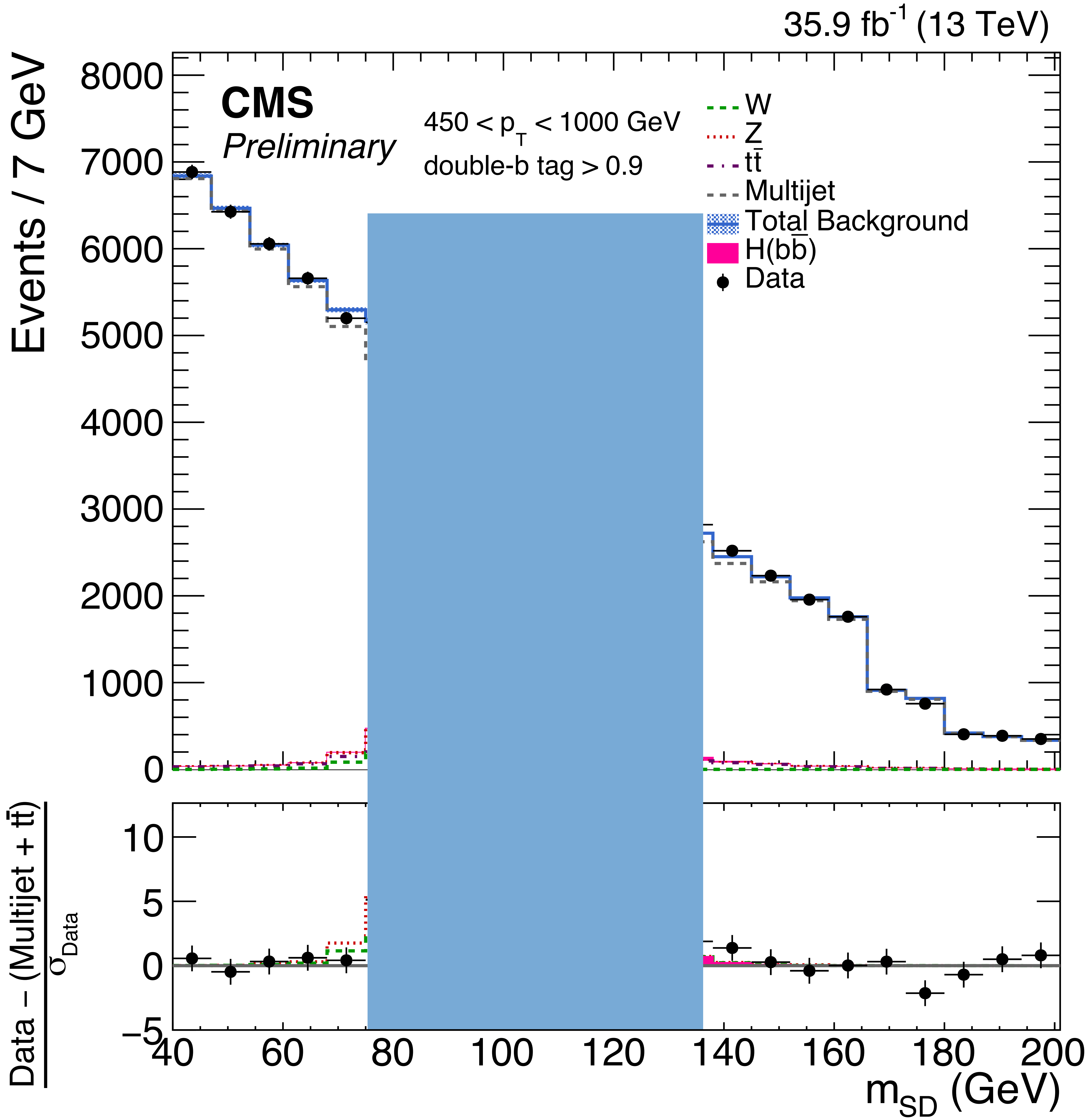
Systematics

Systematic uncertainty source	Type (shape or normalization)	Relative size (or description)
QCD transfer factor	both	profile $a_{k\ell}$ and QCD normalization
Luminosity	normalization	2.5%
V-tag ($N_2^{1,DDT}$) efficiency	normalization	4.3%
Muon veto efficiency	normalization	0.5%
Electron veto efficiency	normalization	0.5%
Trigger efficiency	normalization	4%
Muon ID efficiency	shape	up to 0.2%
Muon isolation efficiency	shape	up to 0.1%
Muon trigger efficiency	shape	up to 8%
$t\bar{t}$ normalization SF	normalization	from 1μ CR: 8%
$t\bar{t}$ double-b mis-tag SF	normalization	from 1μ CR: 15%
W/Z NLO QCD corrections	normalization	10%
W/Z NLO EWK corrections	normalization	15% – 35%
W/Z NLO EWK ratio decorrelation	normalization	5% – 15%
double-b tagging efficiency	normalization	4%
Jet energy scale	normalization	up to 10%
Jet energy resolution	normalization	up to 15%
Jet mass scale	shape	shift m_{SD} peak by $\pm$0.4%
Jet mass resolution	shape	smear m_{SD} distribution by $\pm$9%
Jet mass scale p_T	normalization	0.4%/100 GeV (p_T)
Monte Carlo statistics	normalization	-
H p_T correction (gluon fusion)	both	30%

correlated among W, Z, and H

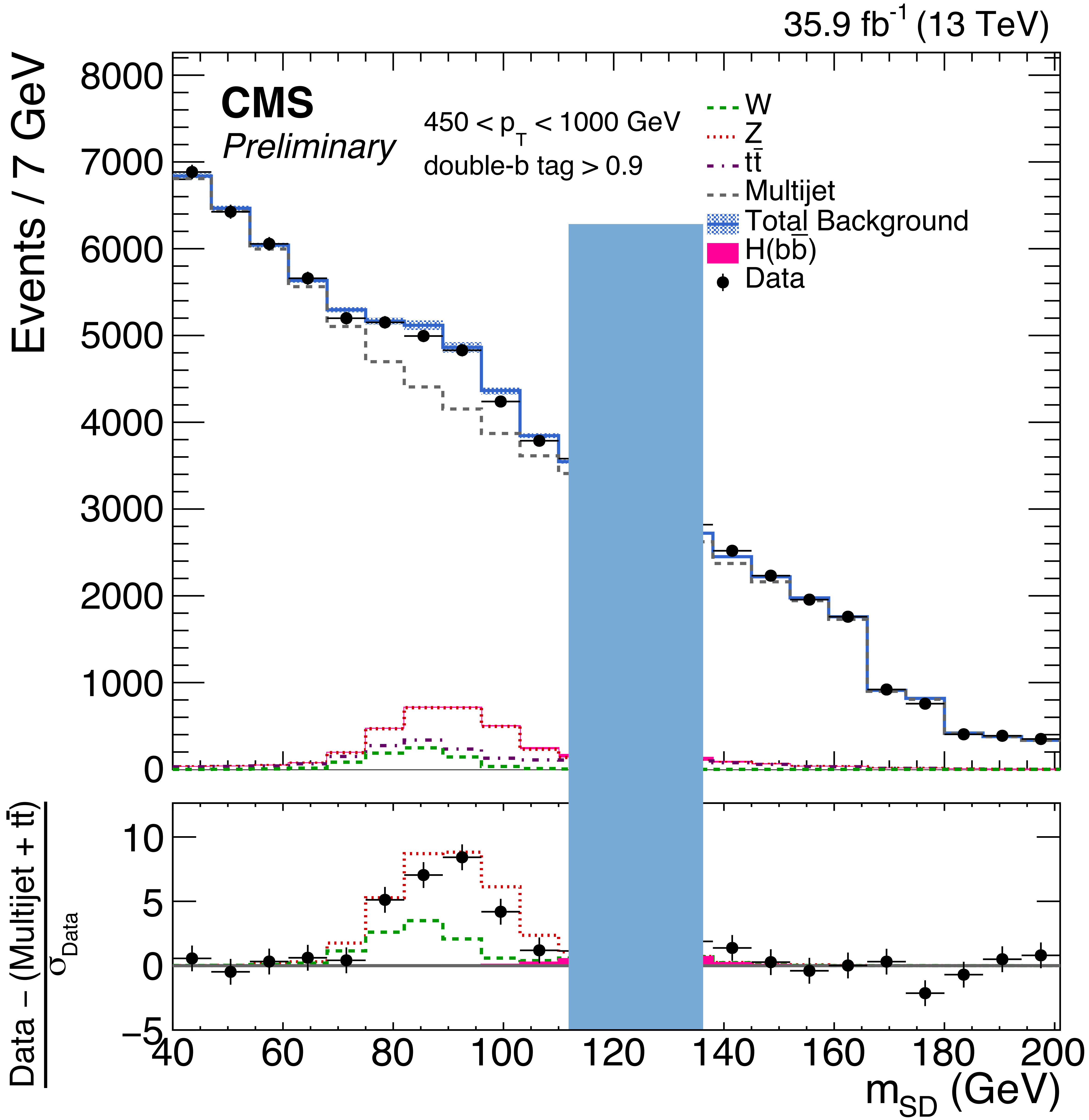
Z signal extraction

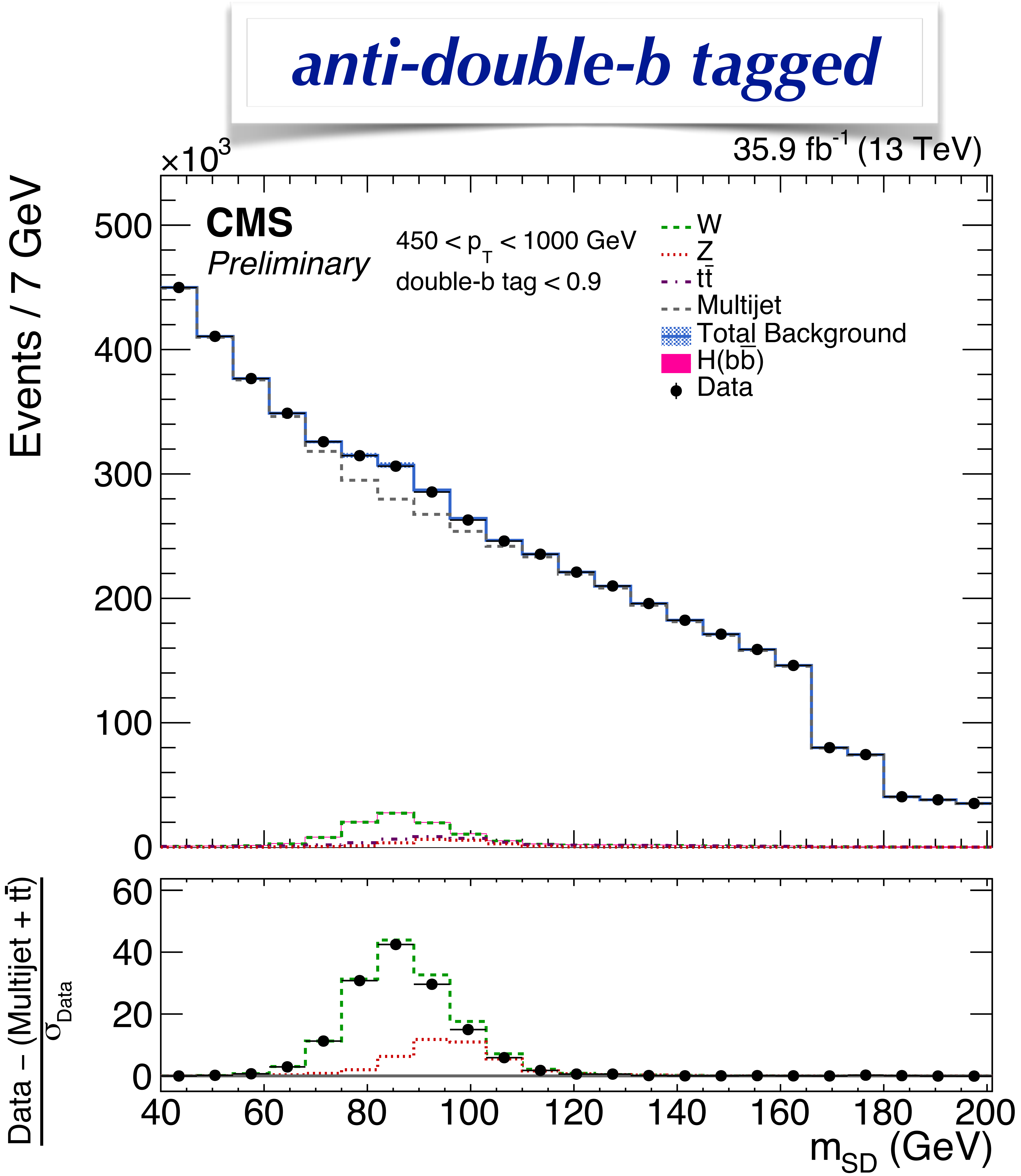
- The extraction of the Z signal compatible with the SM expectation would validate the H signal extraction and H-tagging approach



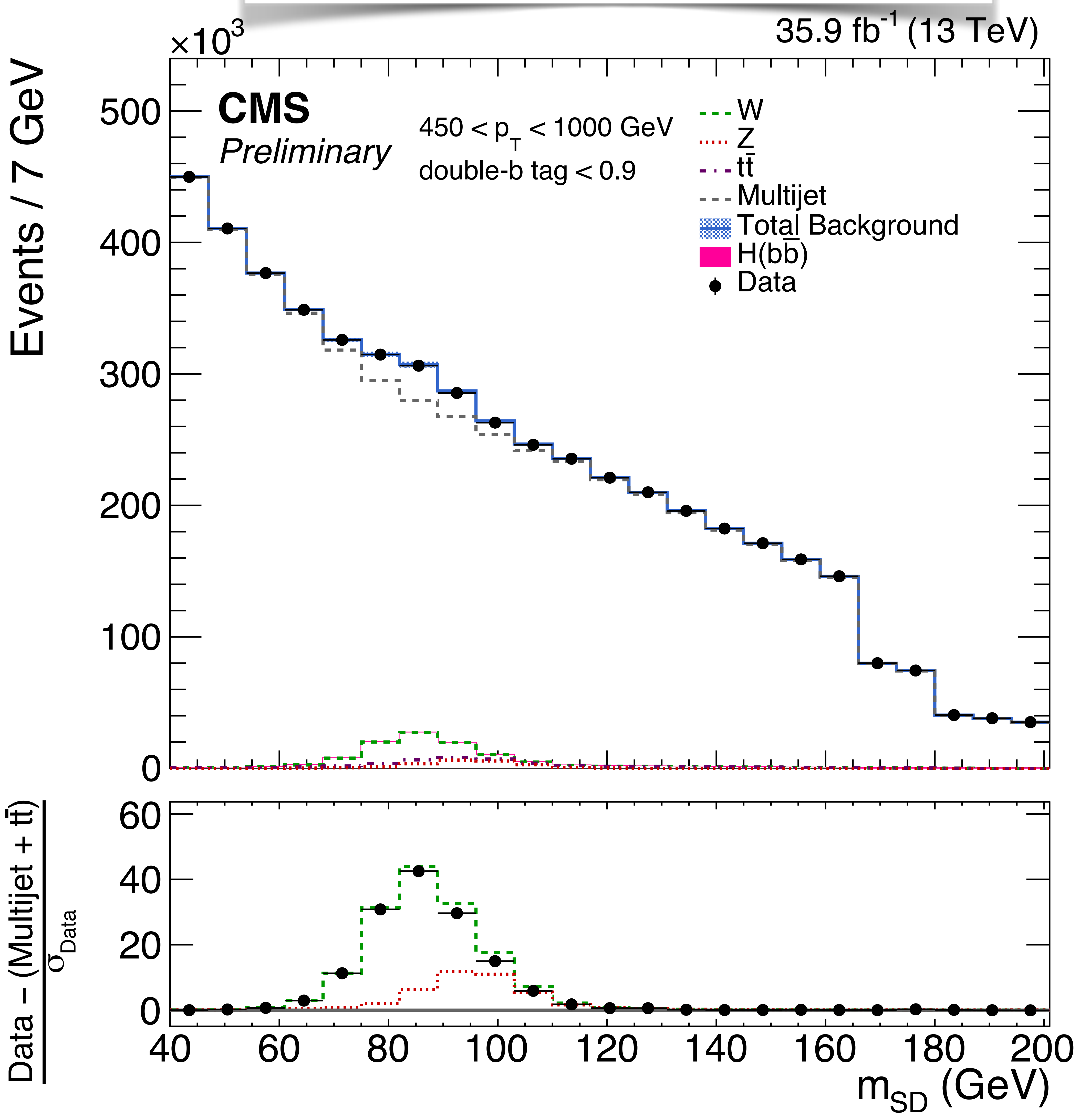
Z results

- Observed significance for the Z signal is 5.1σ (5.8 σ expected)
  - **compatible with SM expectation**
- ***First Observation of the Z(b$\bar{b}$) in the one-jet topology***
- This validates the H signal extraction and H-tagging approach

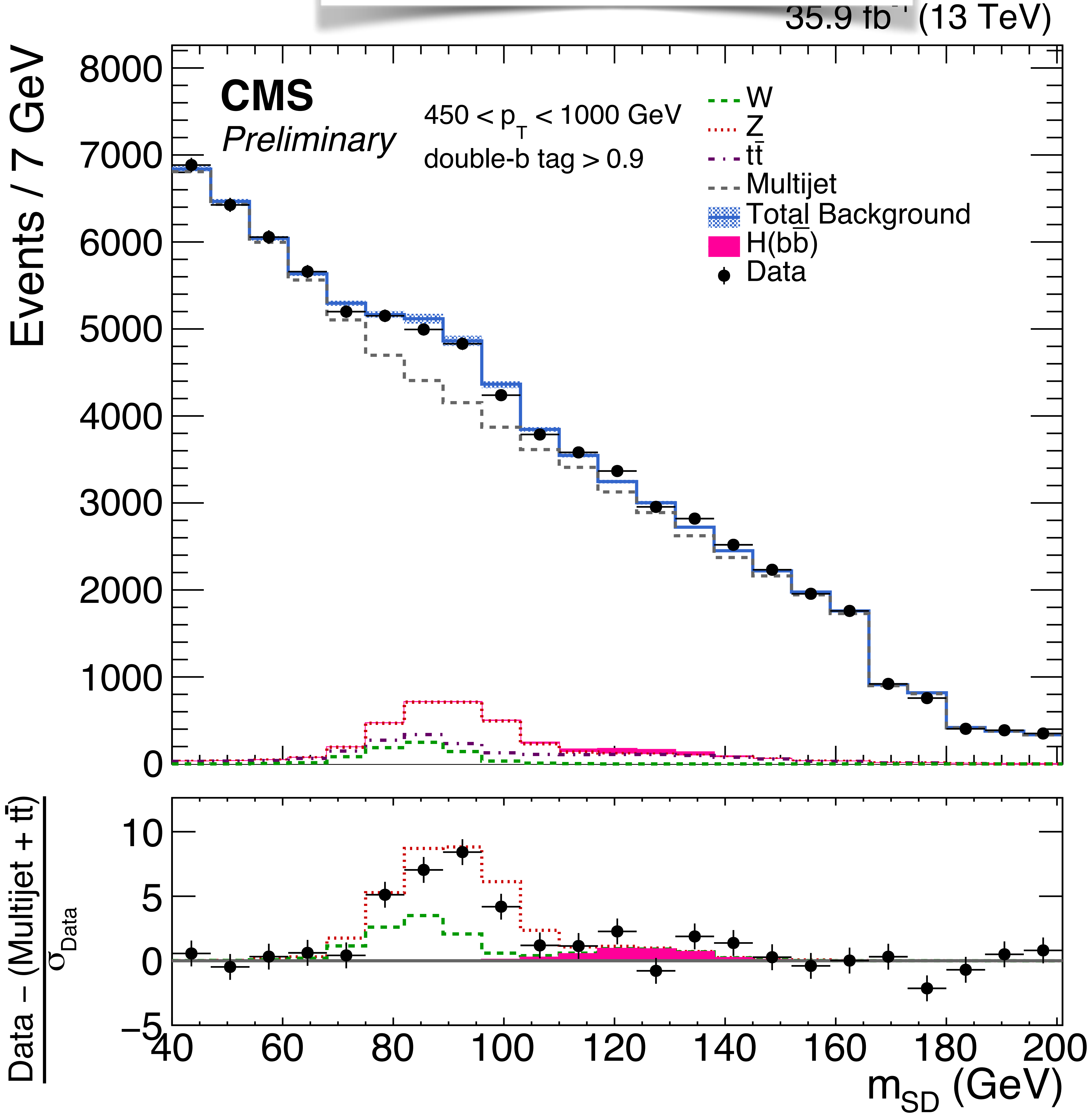



anti-double-b tagged
35.9 fb-1 (13 TeV)
×10^3
CMS
Preliminary
450 < pT < 1000 GeV
double-b tag < 0.9
W
Z
tt̄
Multijet
Total Background
H(bb̄)
Data
Events / 7 GeV
Data − (Multijet + tt̄) / σData
mSD (GeV)

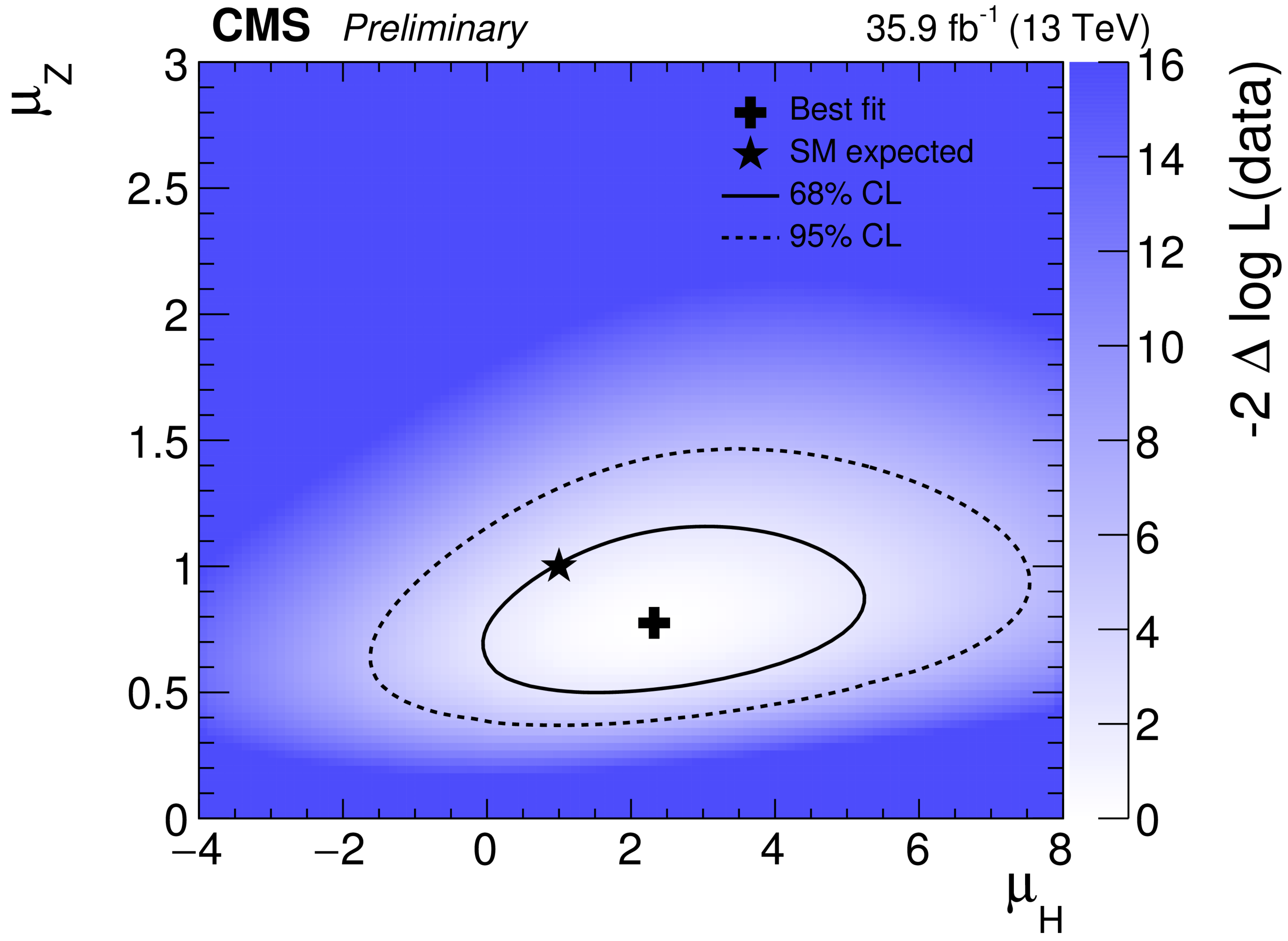
anti-double-b tagged



double-b tagged

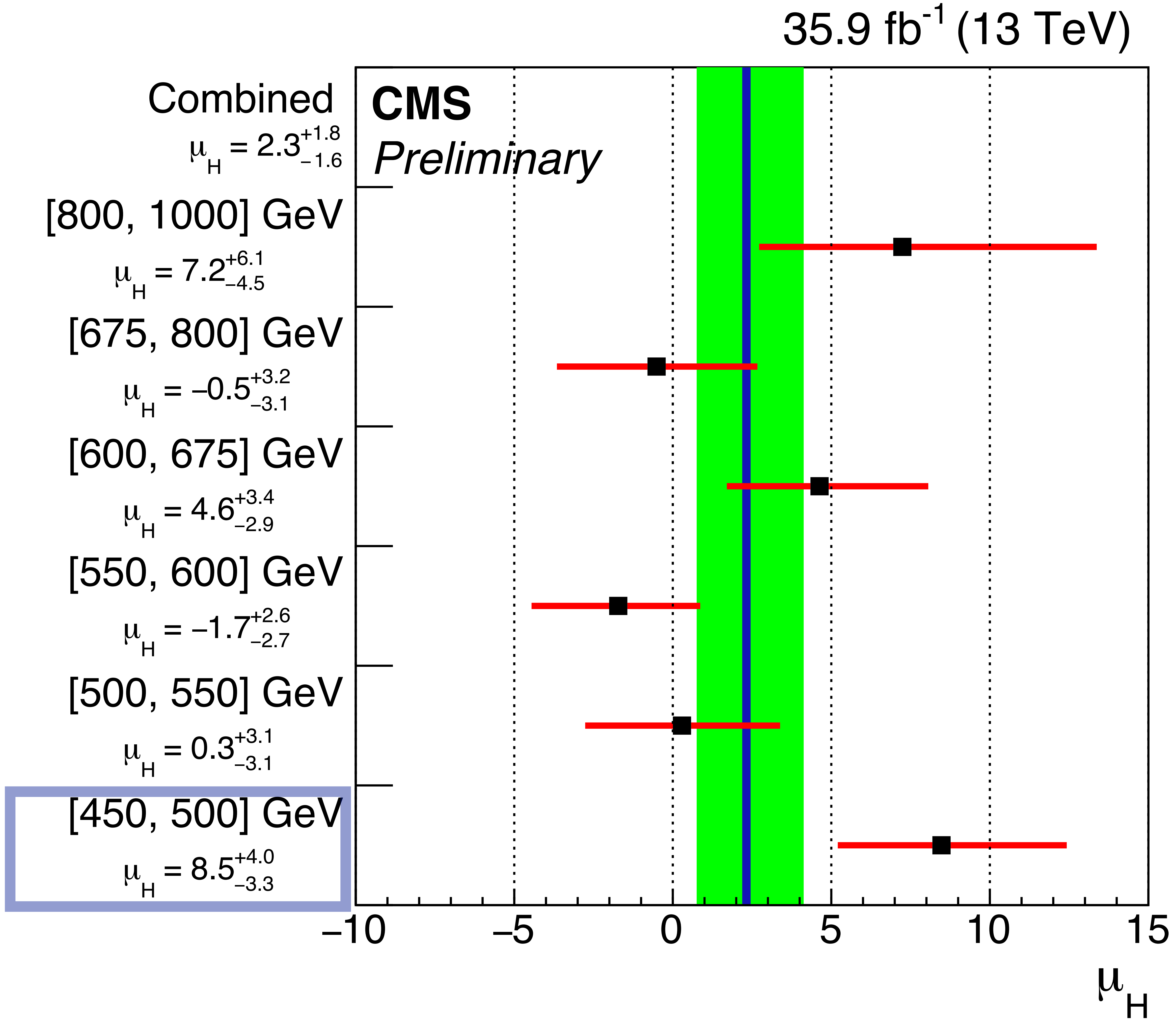


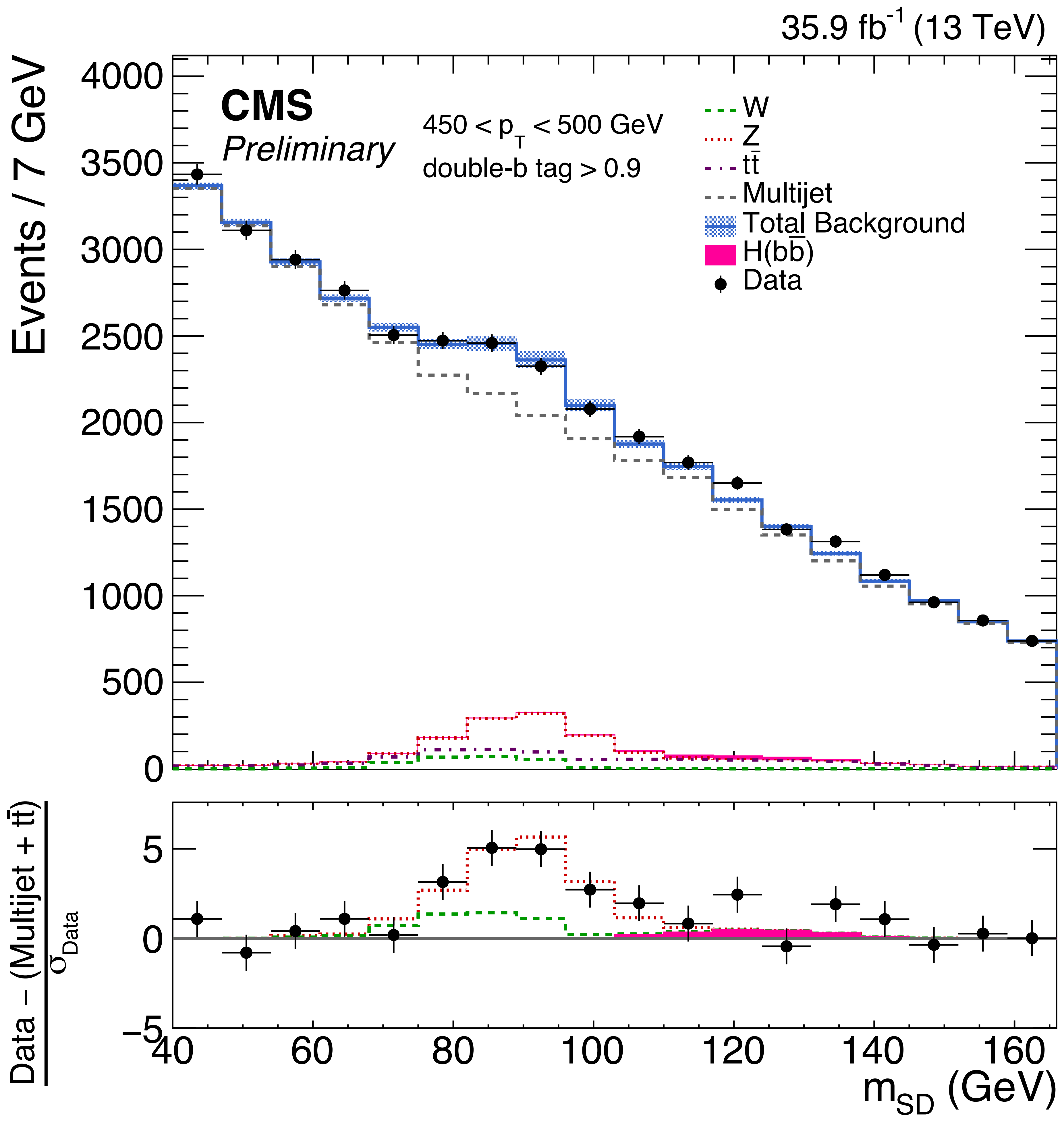
Simultaneous fit of the Z and H signals



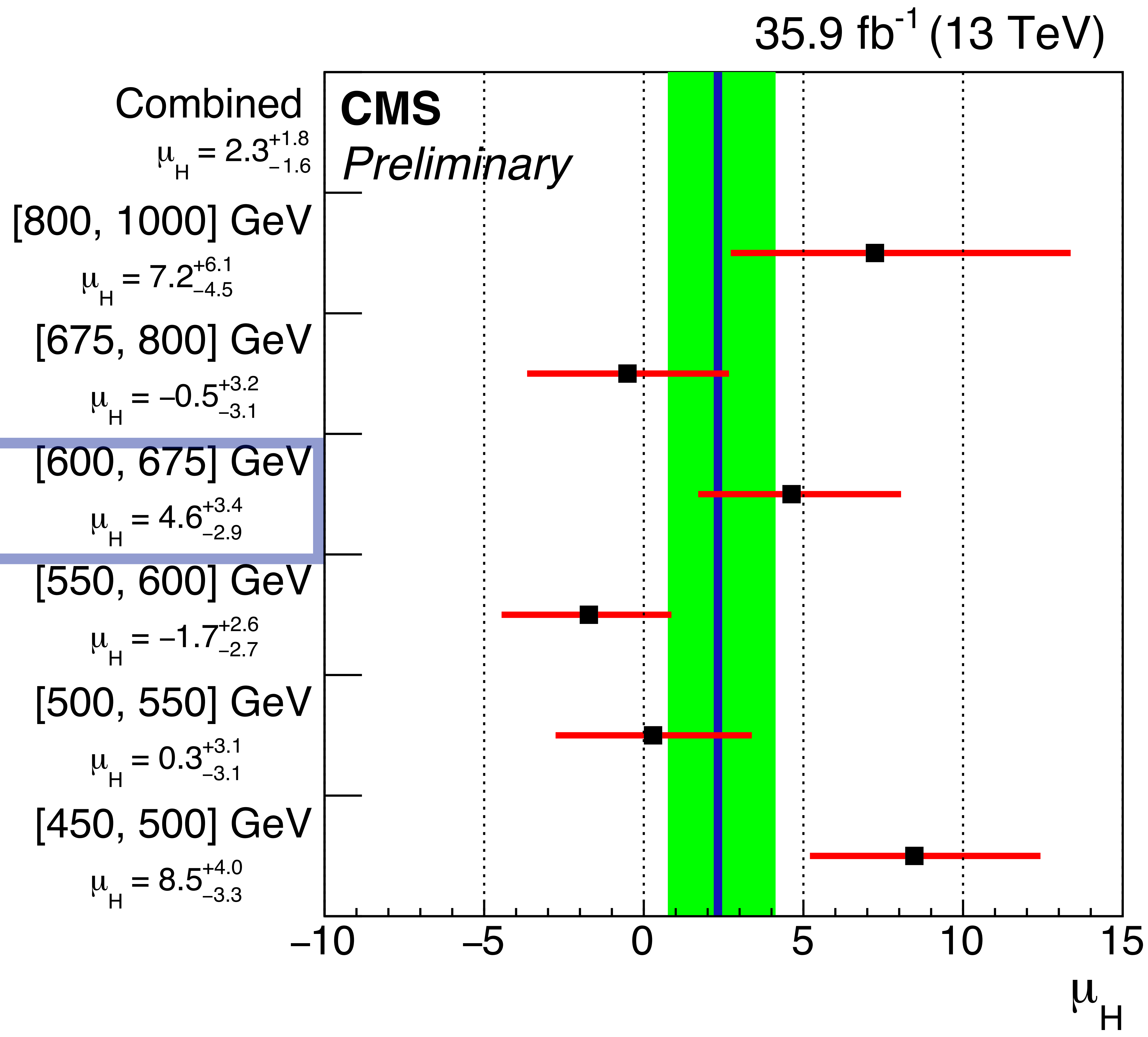
	H	Z
Observed best fit	$\mu_H = 2.3^{+1.8}_{-1.6}$	$\mu_Z = 0.78^{+0.23}_{-0.19}$
Expected significance	0.7σ ($\mu_H = 1$)	5.8σ ($\mu_Z = 1$)
Observed significance	1.5σ	5.1σ

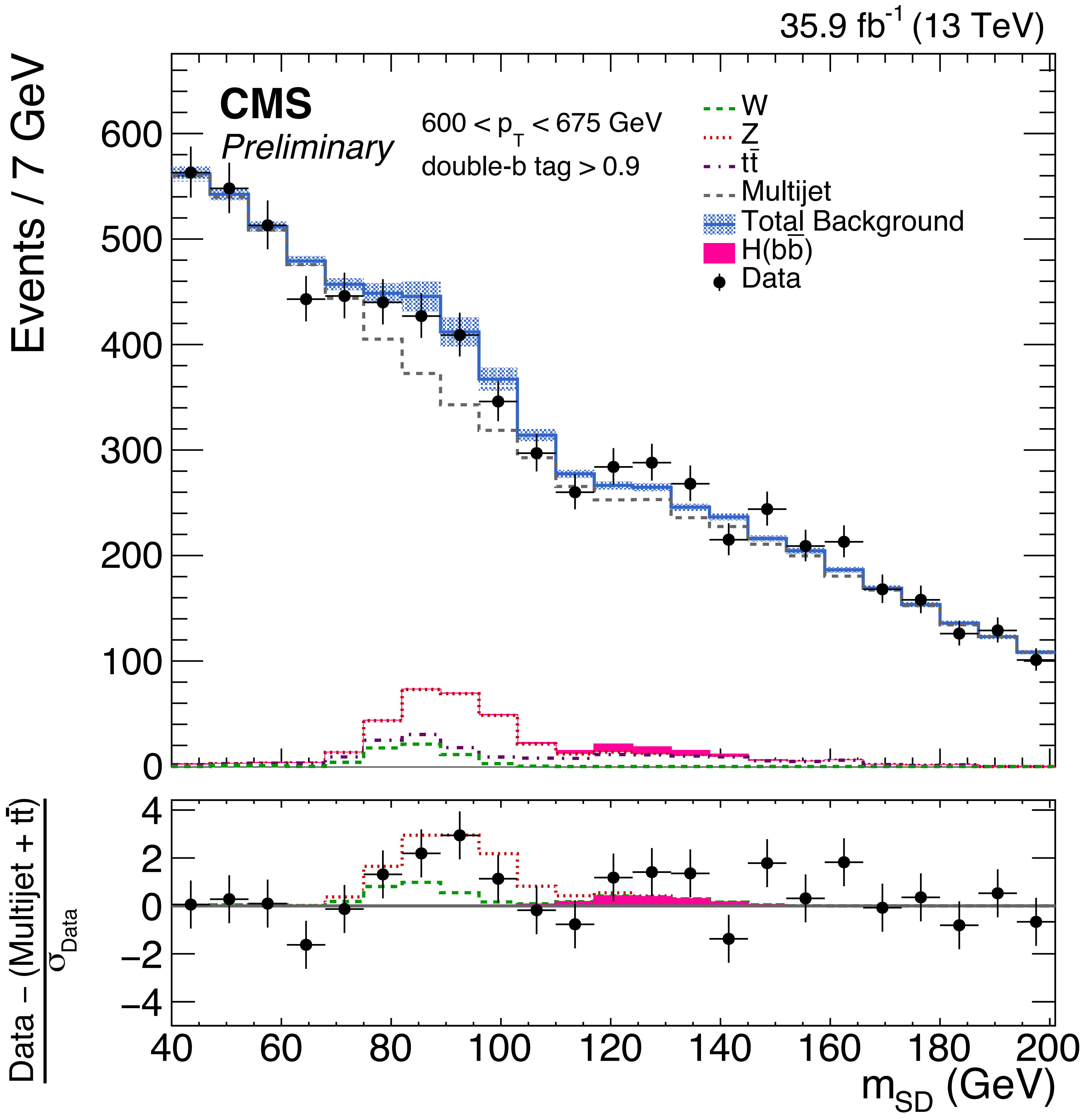
Sensitivity per p_T category



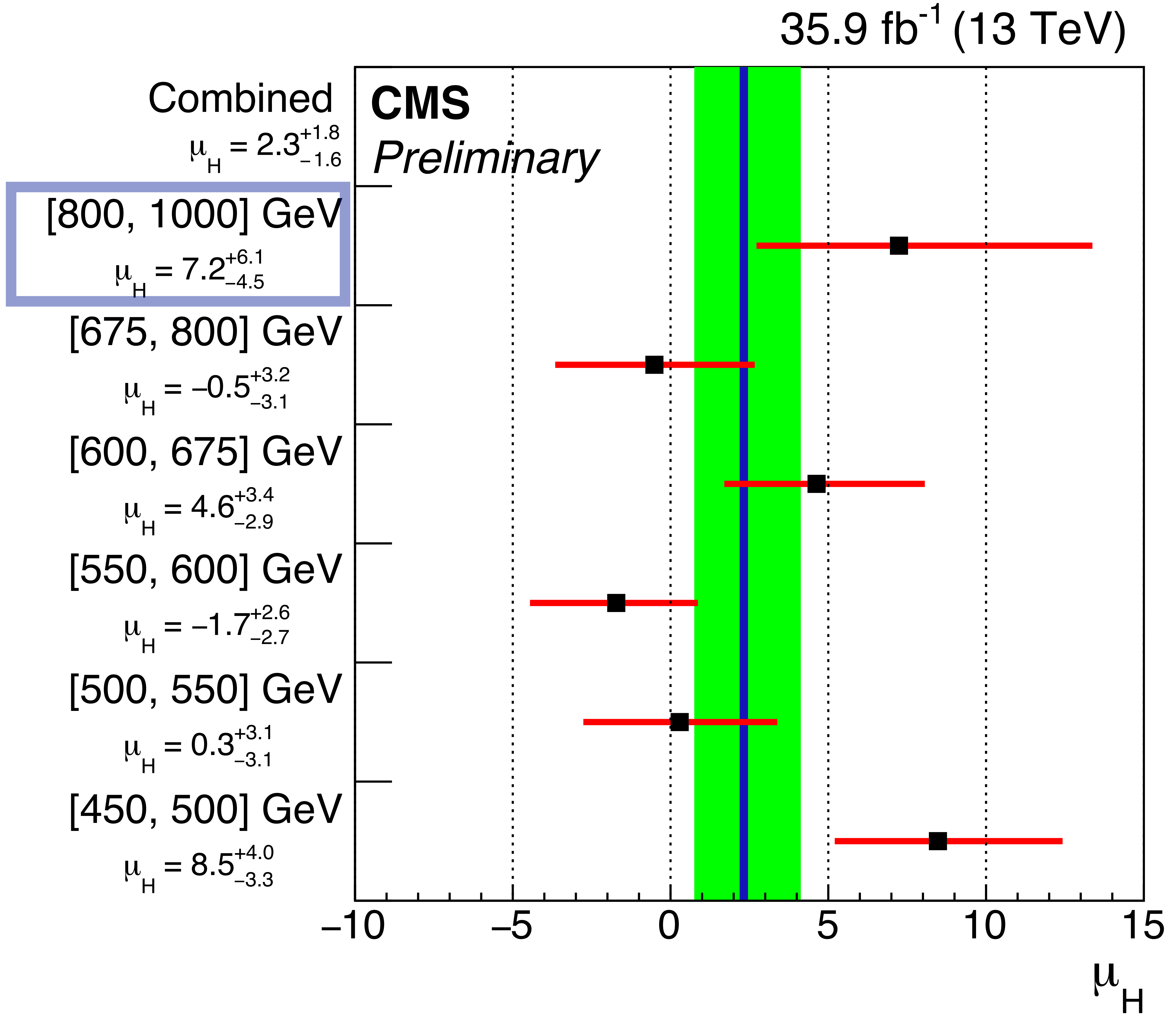


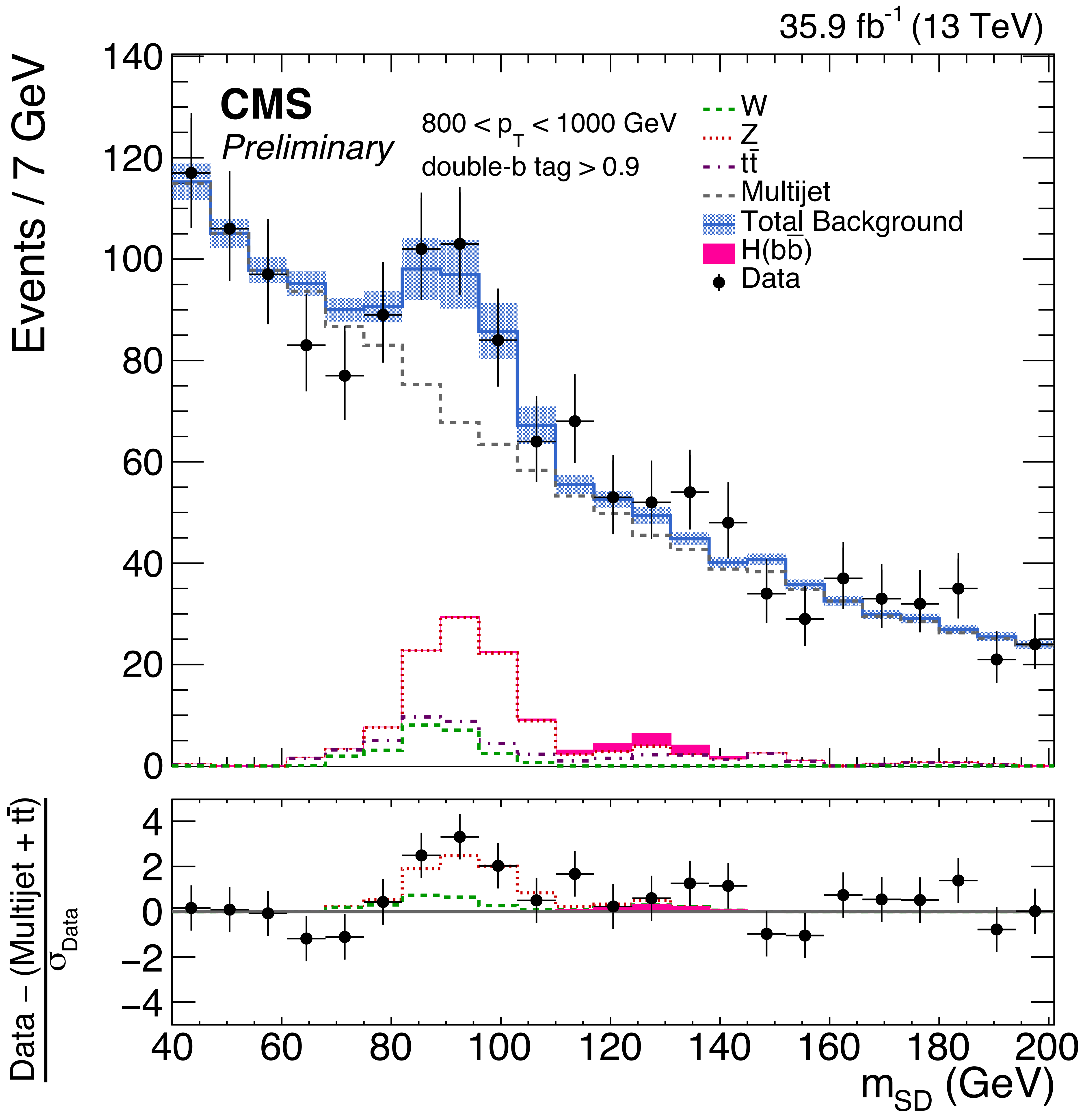
Sensitivity per p_T category





Sensitivity per p_T category





Measured cross section

- The measured cross sections for Z+jets and Higgs for jet p_T > 450 GeV are:

$$\sigma_Z = 849 \; {+257}/{-209} \text{ fb}$$
$$\sigma_H = 74 \; {+51}/{-49} \text{ fb}$$

- Broken down into:

$$\sigma_Z = 849 \; {+155}/{-155} \text{ (stat.) } {+140}/{-205} \text{ (syst.)}$$
$$\sigma_H = 74 \; {+48}/{-48} \text{ (stat.) } {+10}/{-17} \text{ (syst.)}$$

Compatible with SM within uncertainties

Conclusions and perspectives

- First LHC search for $gg \to H \to b\bar{b}$ in boosted topology
  - **First observation of $Z(b\bar{b})$** in single jet topology, 5.1σ (5.8 σ expected)
  - The observed significance for the $H(b\bar{b})$ is 1.5σ
  - Cross sections are measured and agree with SM

$$\sigma_Z = 849 \ {+155/-155} \ (\text{stat.}) \ {+140/-205} \ (\text{syst.})$$
$$\sigma_H = 74 \ {+48/-48} \ (\text{stat.}) \ {+10/-17} \ (\text{syst.})$$

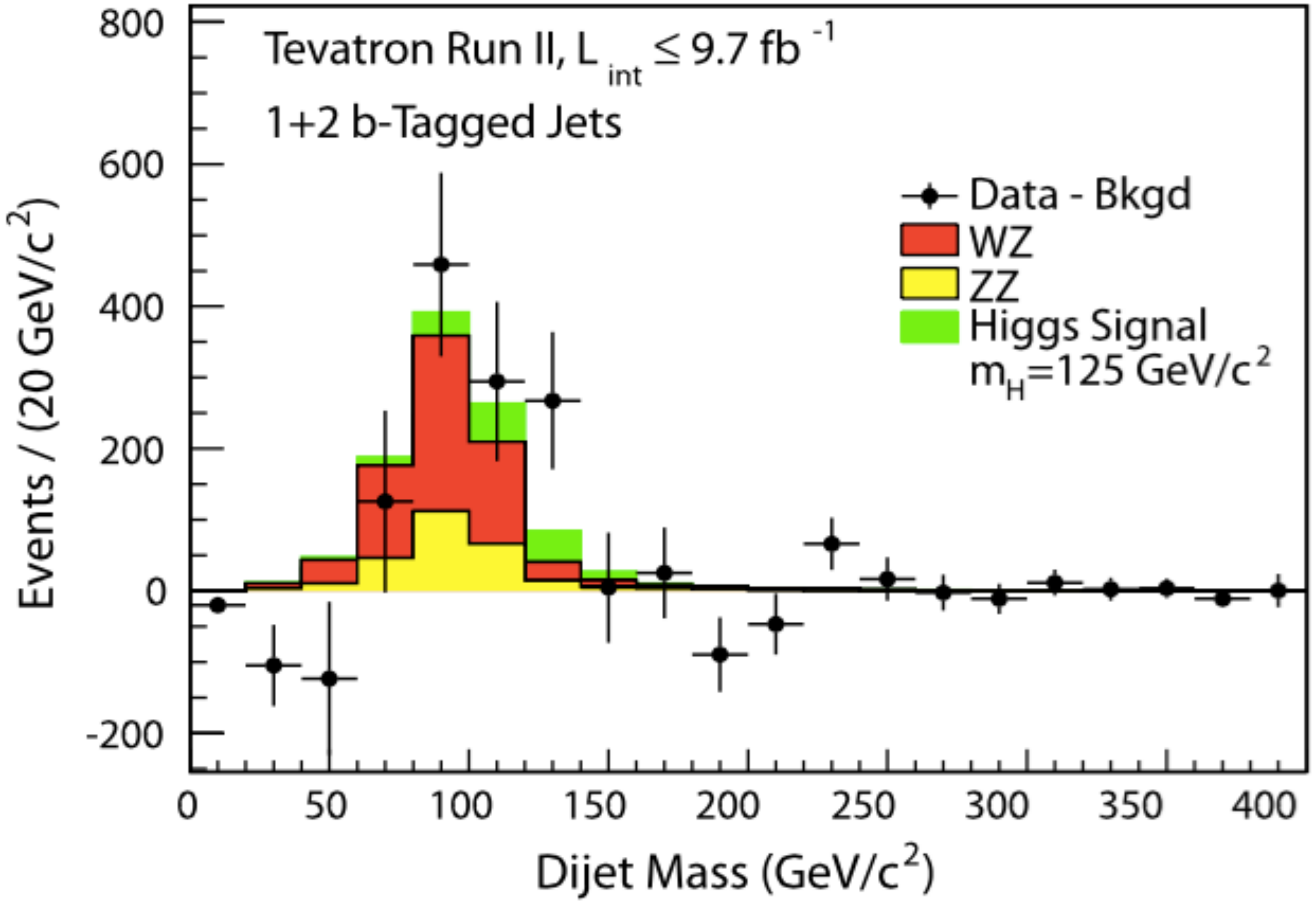
- This search probes previously unexplored regions of phase space
  - open a **new strategy** to search for Higgs boson to $b\bar{b}$
  - but also change to probe **unexplored new physics contributions to the Higgs at very high p_T**

–Additional Material

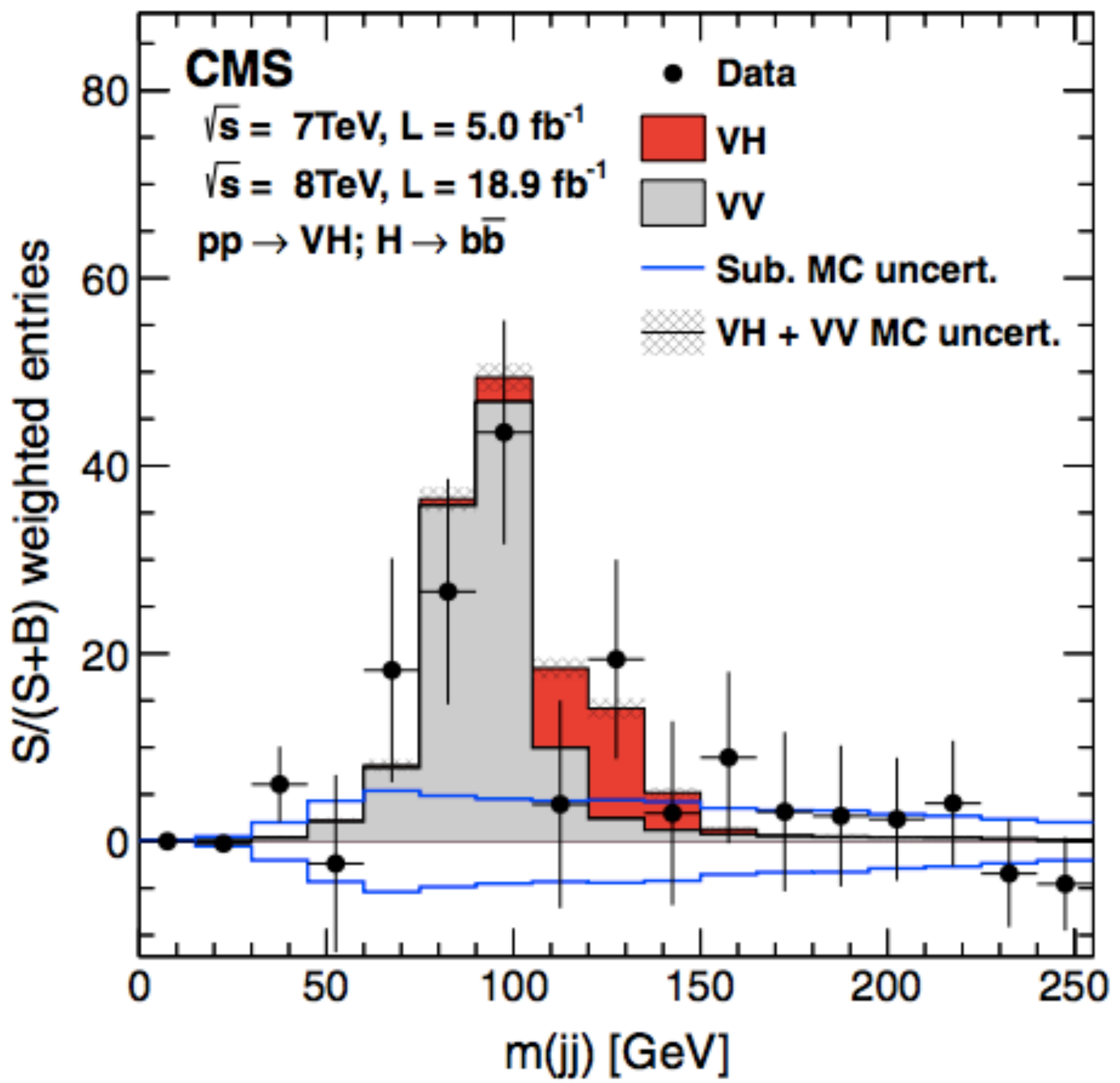
VH to $b\bar{b}$

Phys. Rev. D 89 (2014)
Phys. Rev. Lett. 109 071804

CDF+D0, 2.8σ



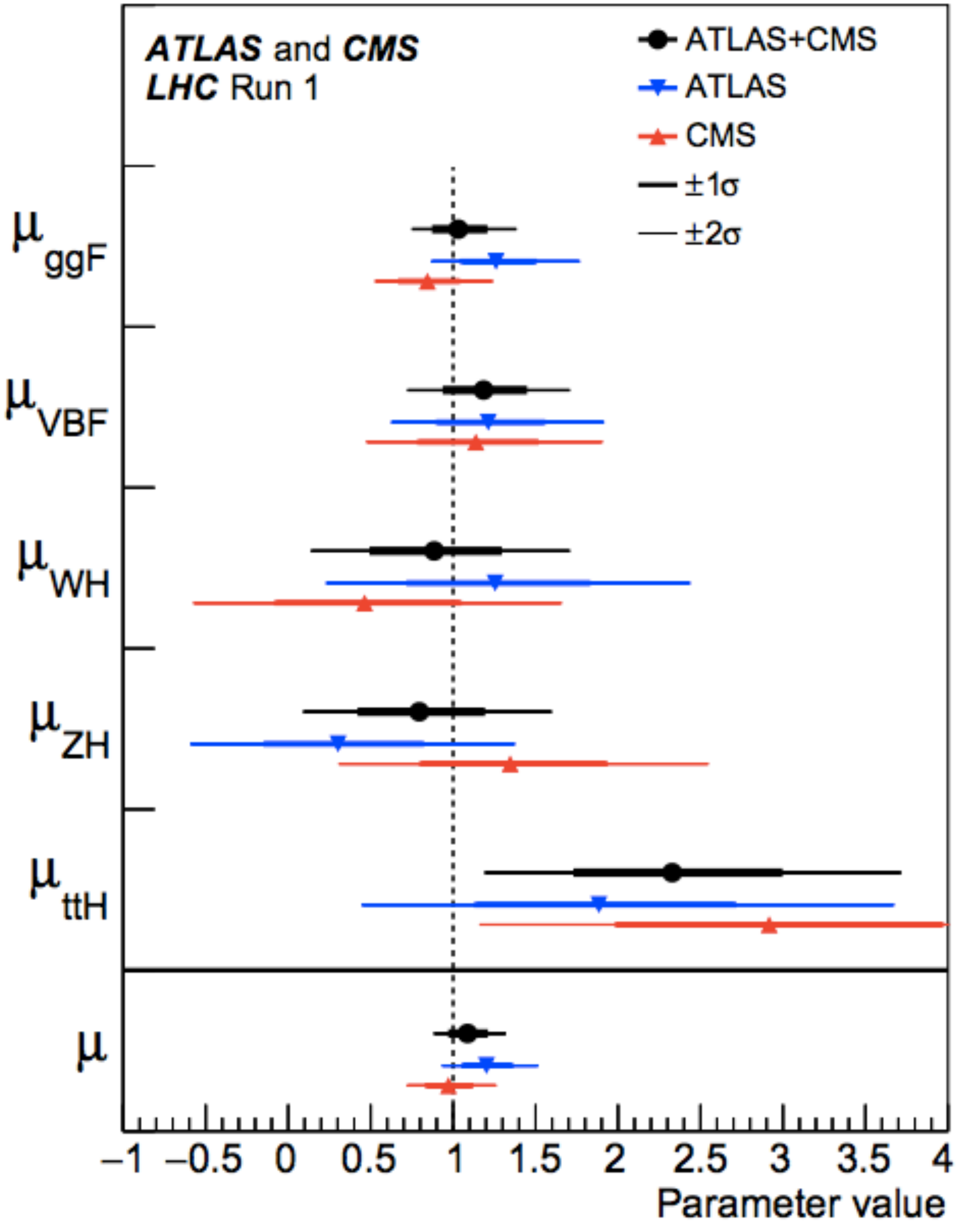
CMS 2.1σ

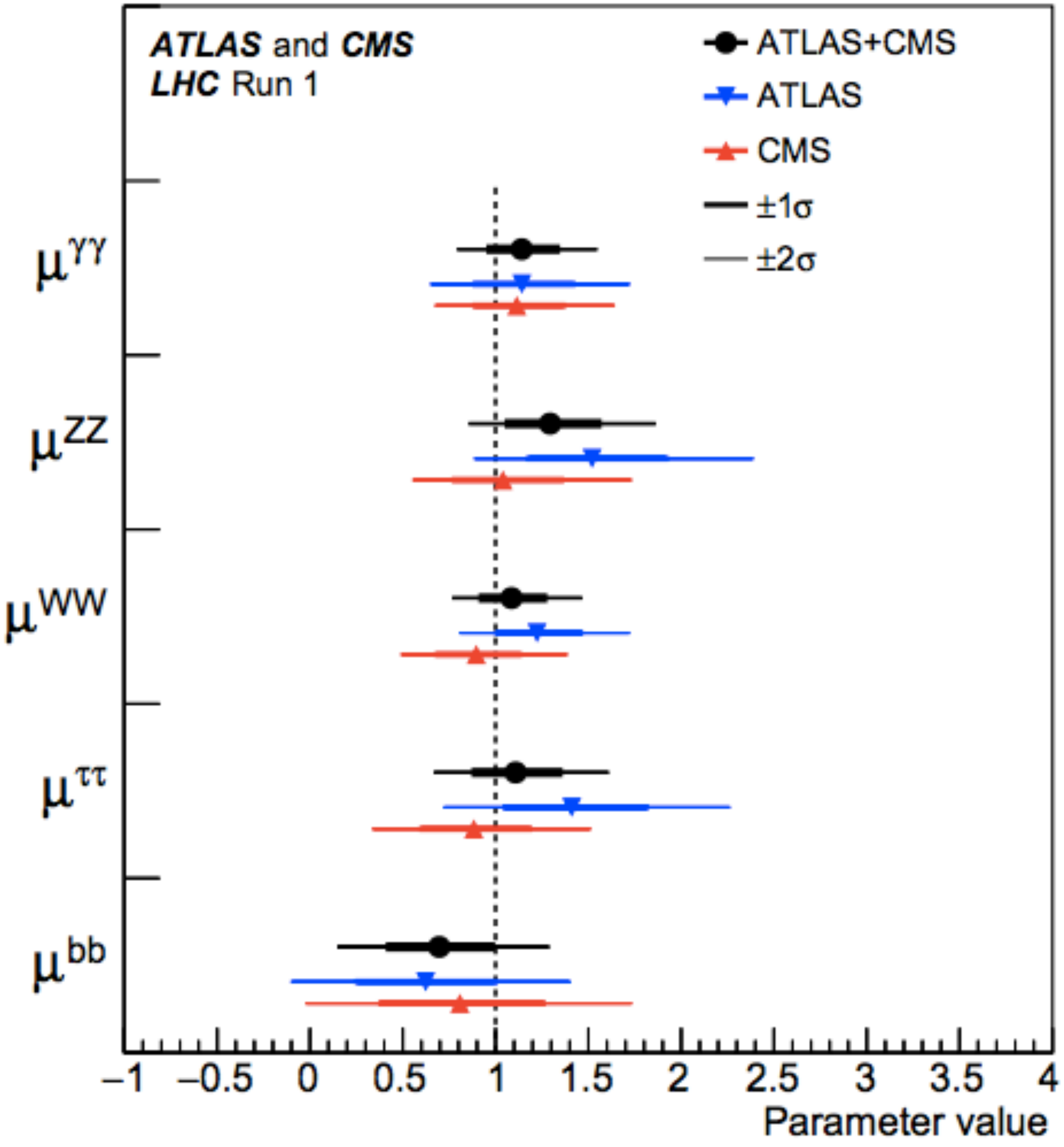


S/B at LHC is 2.5 lower than at Tevatron

Results from ATLAS and CMS with full 13 TeV dataset are on their way

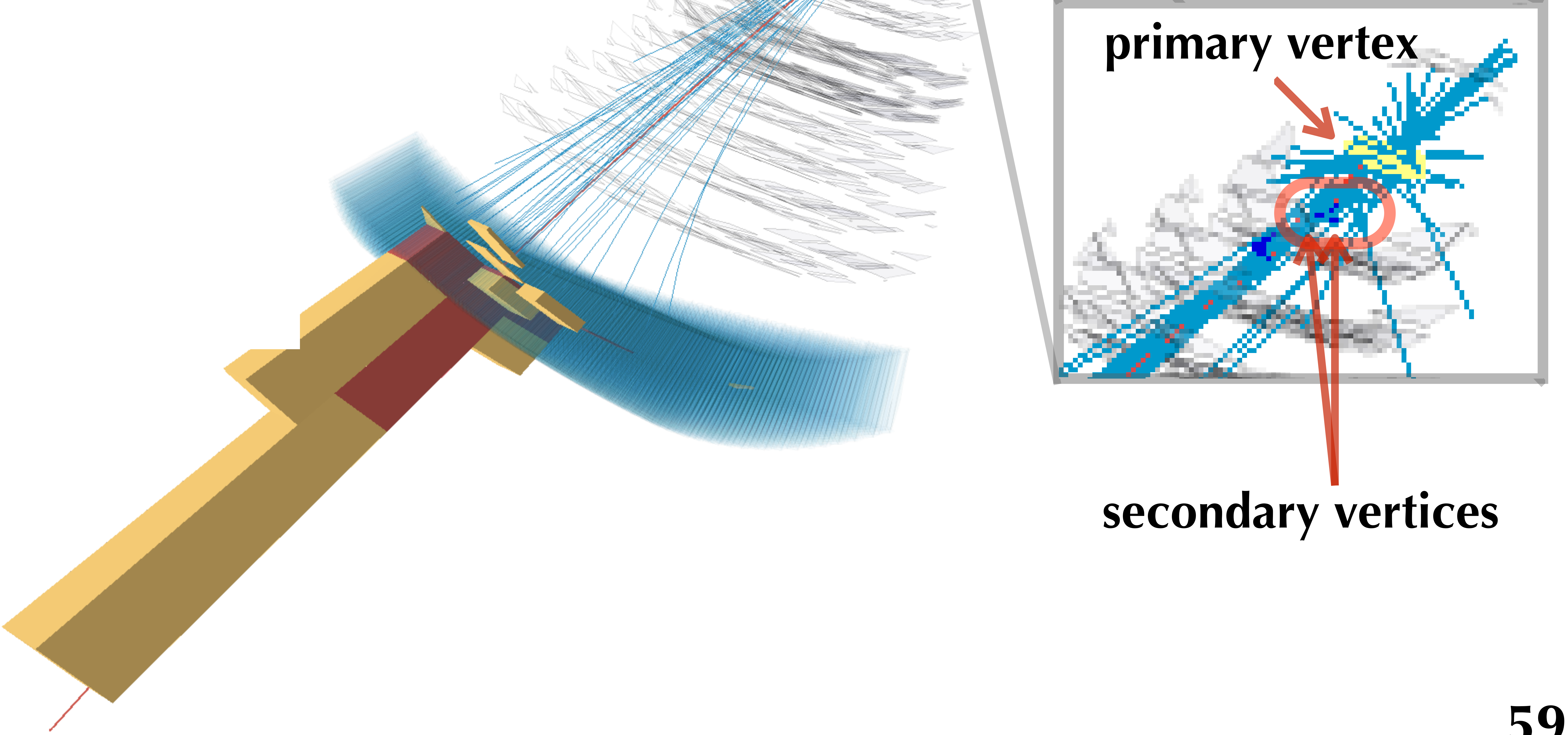
Higgs at LHC


ATLAS and CMS
LHC Run 1
ATLAS+CMS
ATLAS
CMS
±1σ
±2σ
μ_{ggF}
μ_{VBF}
μ_{WH}
μ_{ZH}
μ_{ttH}
μ
−1 −0.5 0 0.5 1 1.5 2 2.5 3 3.5 4
Parameter value


ATLAS and CMS
LHC Run 1
ATLAS+CMS
ATLAS
CMS
±1σ
±2σ
$\mu^{\gamma\gamma}$
μ^{ZZ}
μ^{WW}
$\mu^{\tau\tau}$
μ^{bb}
−1 −0.5 0 0.5 1 1.5 2 2.5 3 3.5 4
Parameter value

$H(b\bar{b})$-jet

p_T	2065 GeV
η	0.63
ϕ	0.84
m_j	123.7 GeV
b tag	0.95



Combined algorithms

The **Combined Secondary Vertex** through multivariate technique combines (CSVv2)

Track information

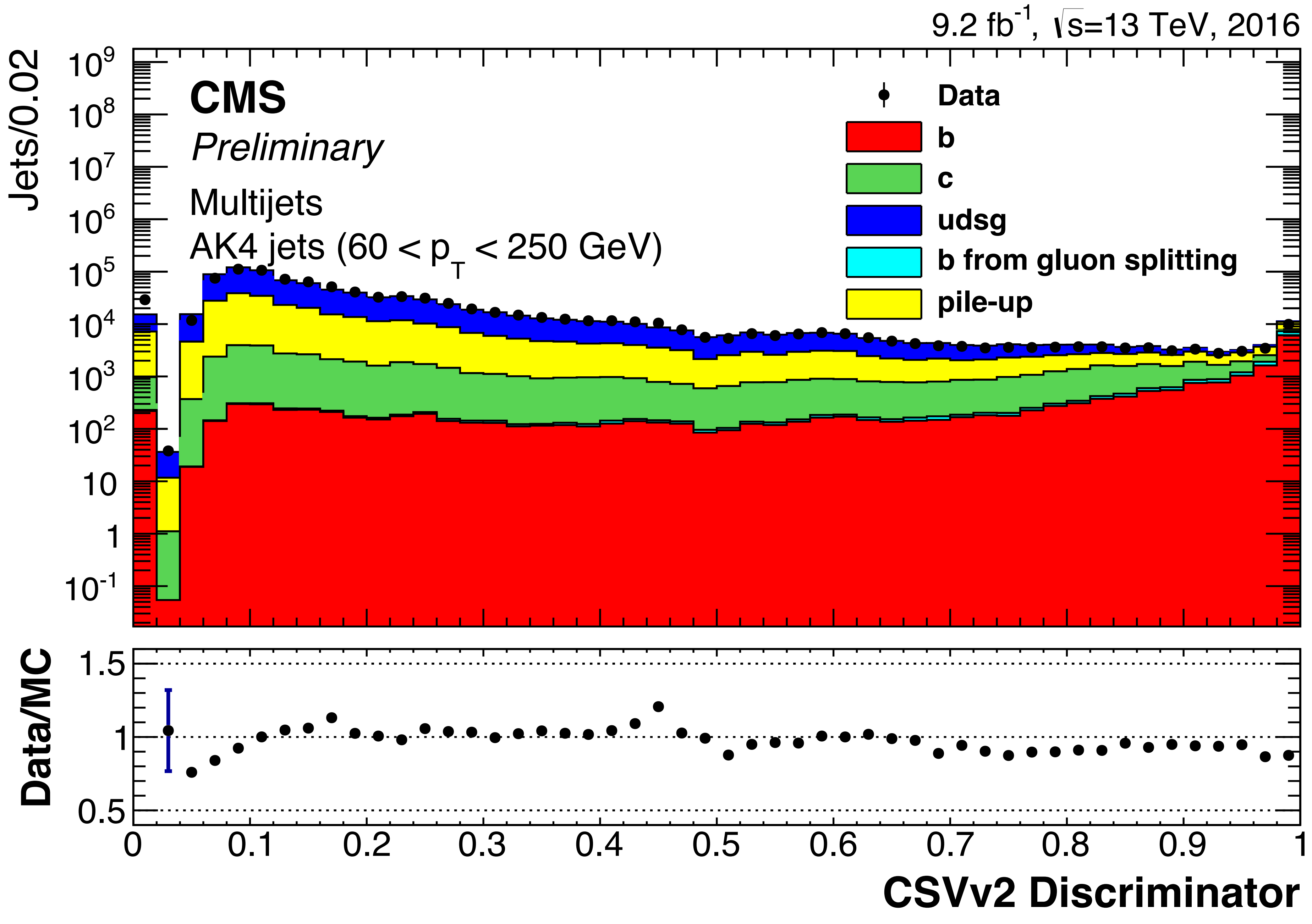
- 3D IP significance of the most energetic tracks

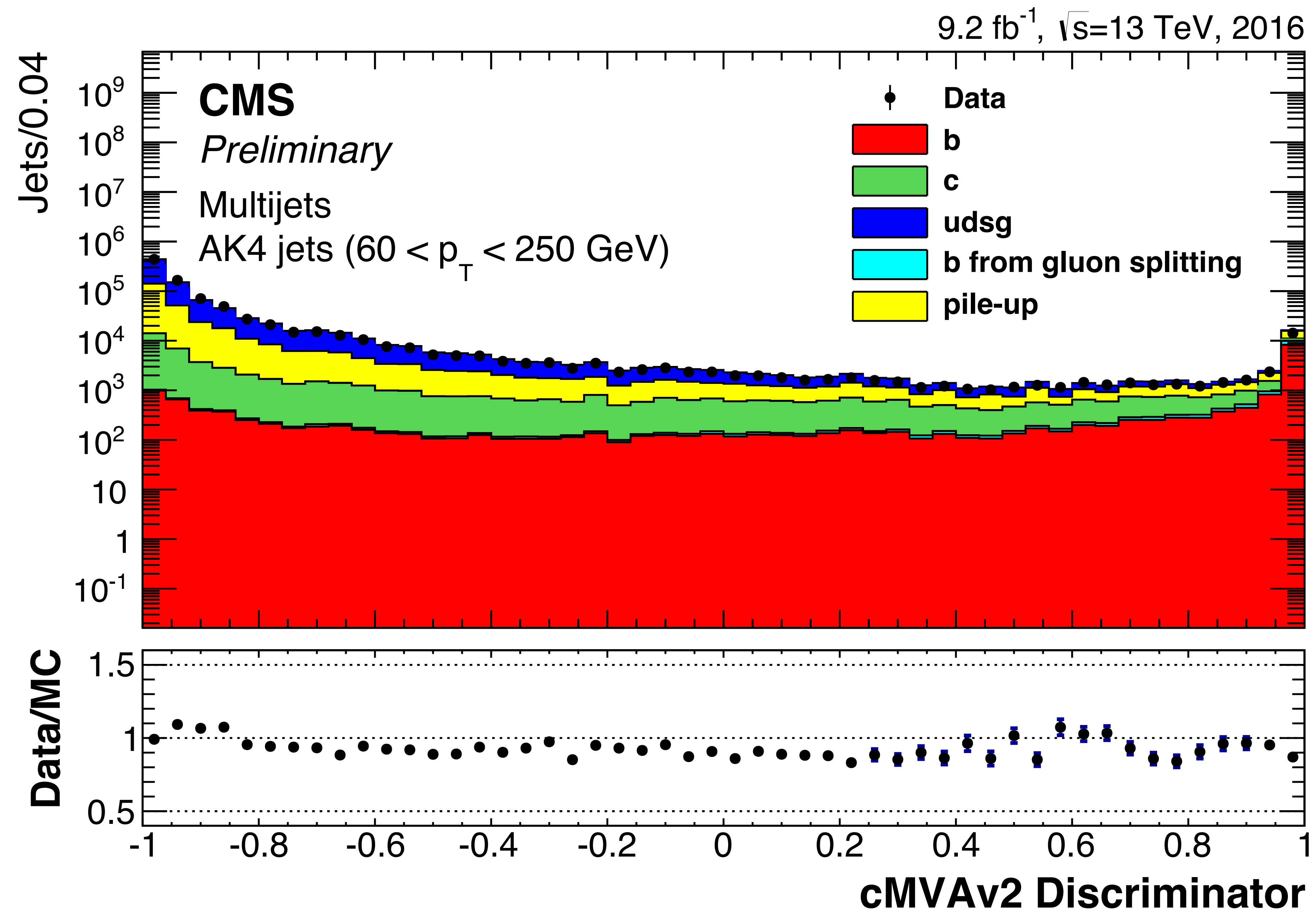
Vertex information

The **Combined Multivariate Algorithm** (cMVAv2) algorithm combines:

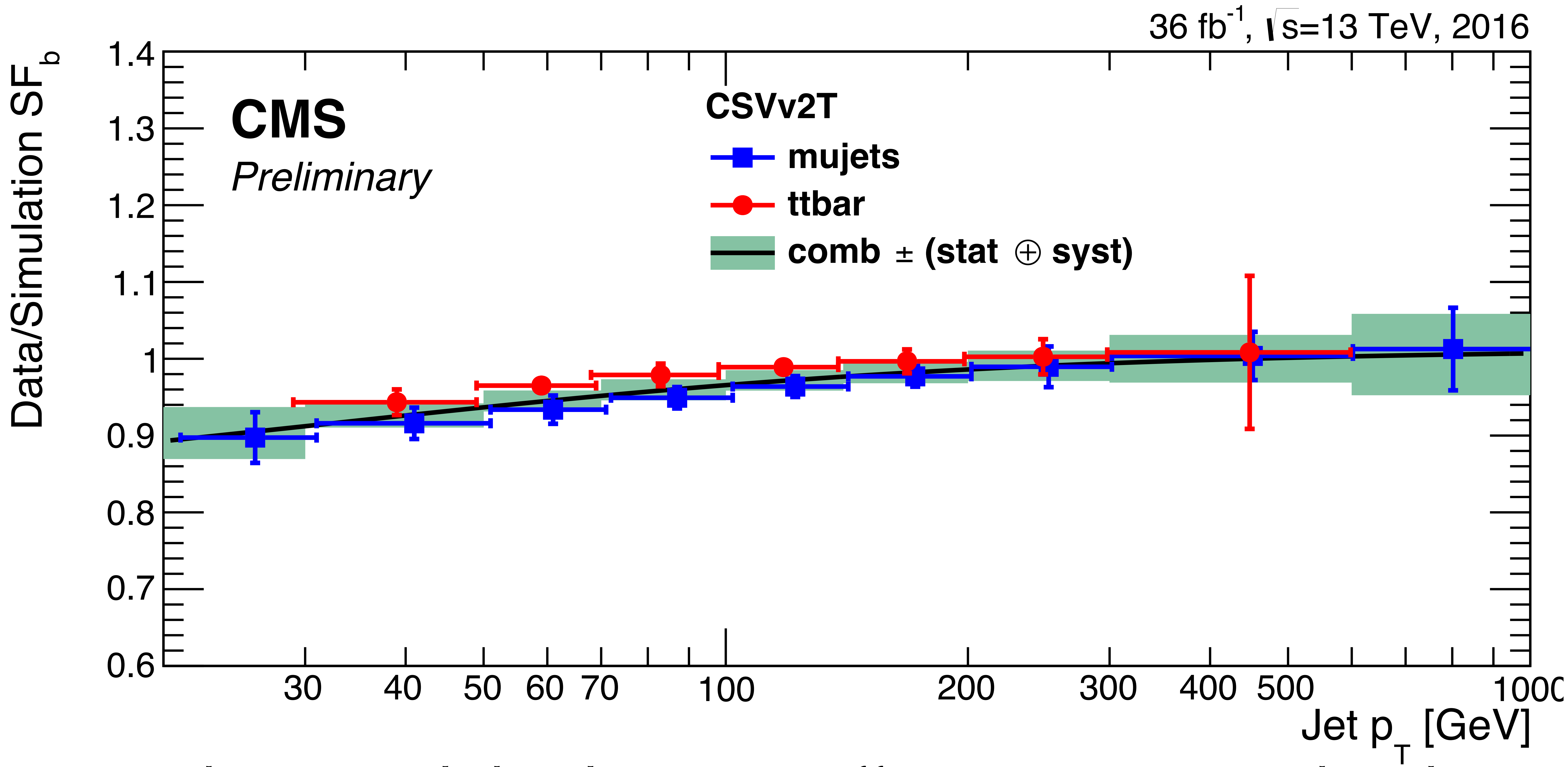
- CSVv2 and soft lepton taggers

DP-16-042





Muon Tagged jets

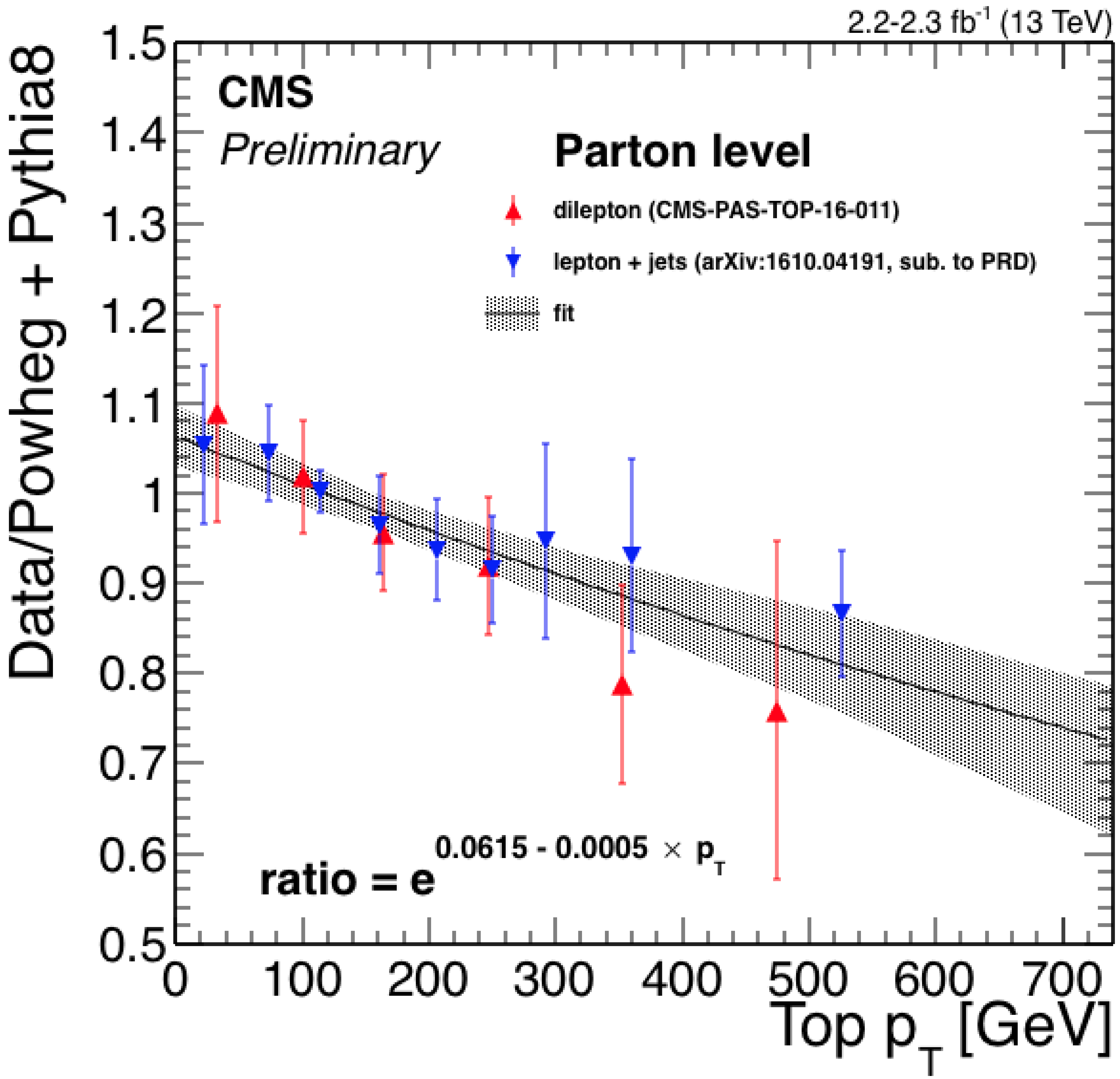


Good compatibility between efficiency measured with:

- muon tagged jets from multijet events
- b-jets from ttbar

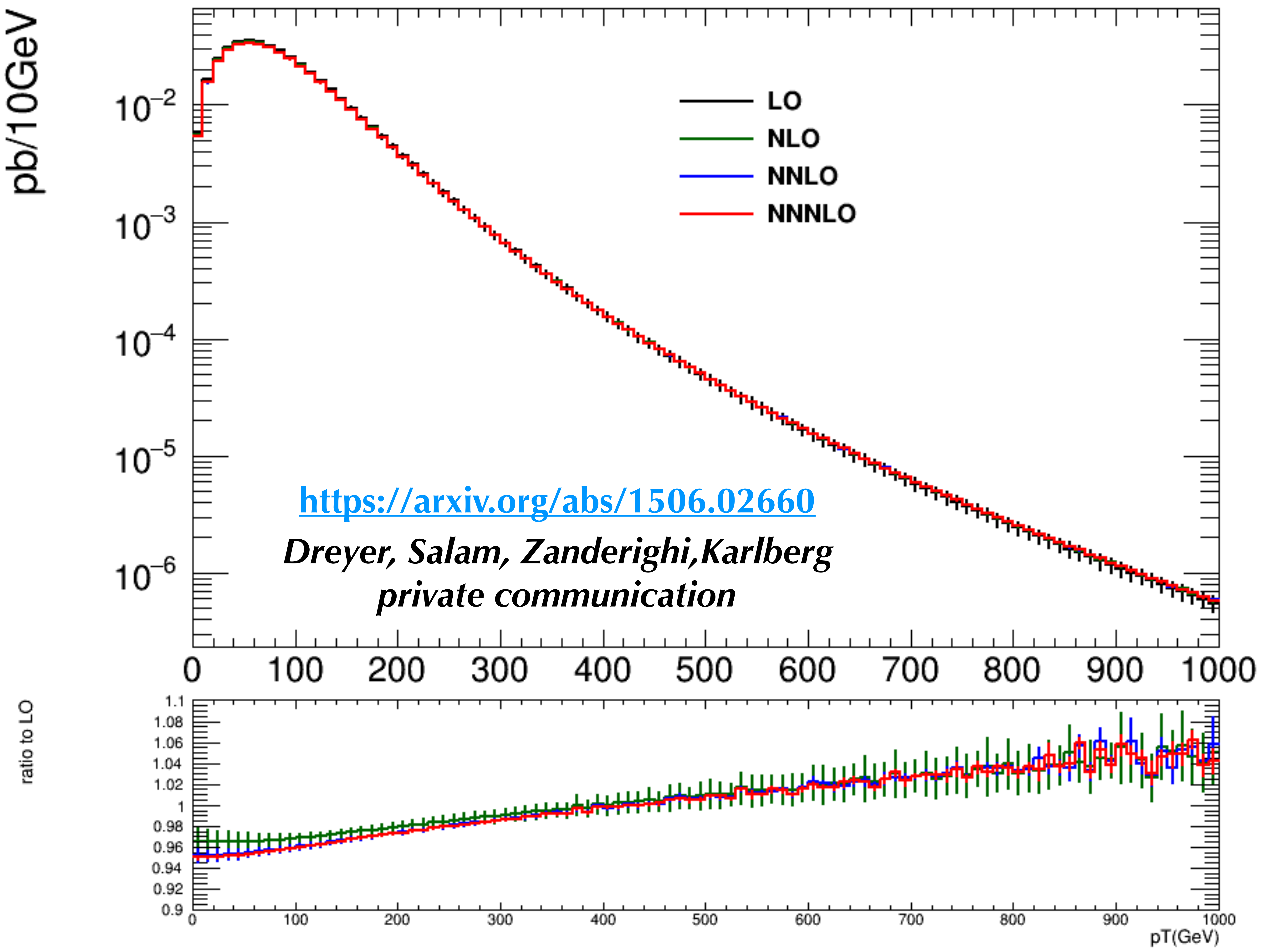
Top p_T reweighting

- Applying 13 TeV top p_T reweighting



VBF H reweighting

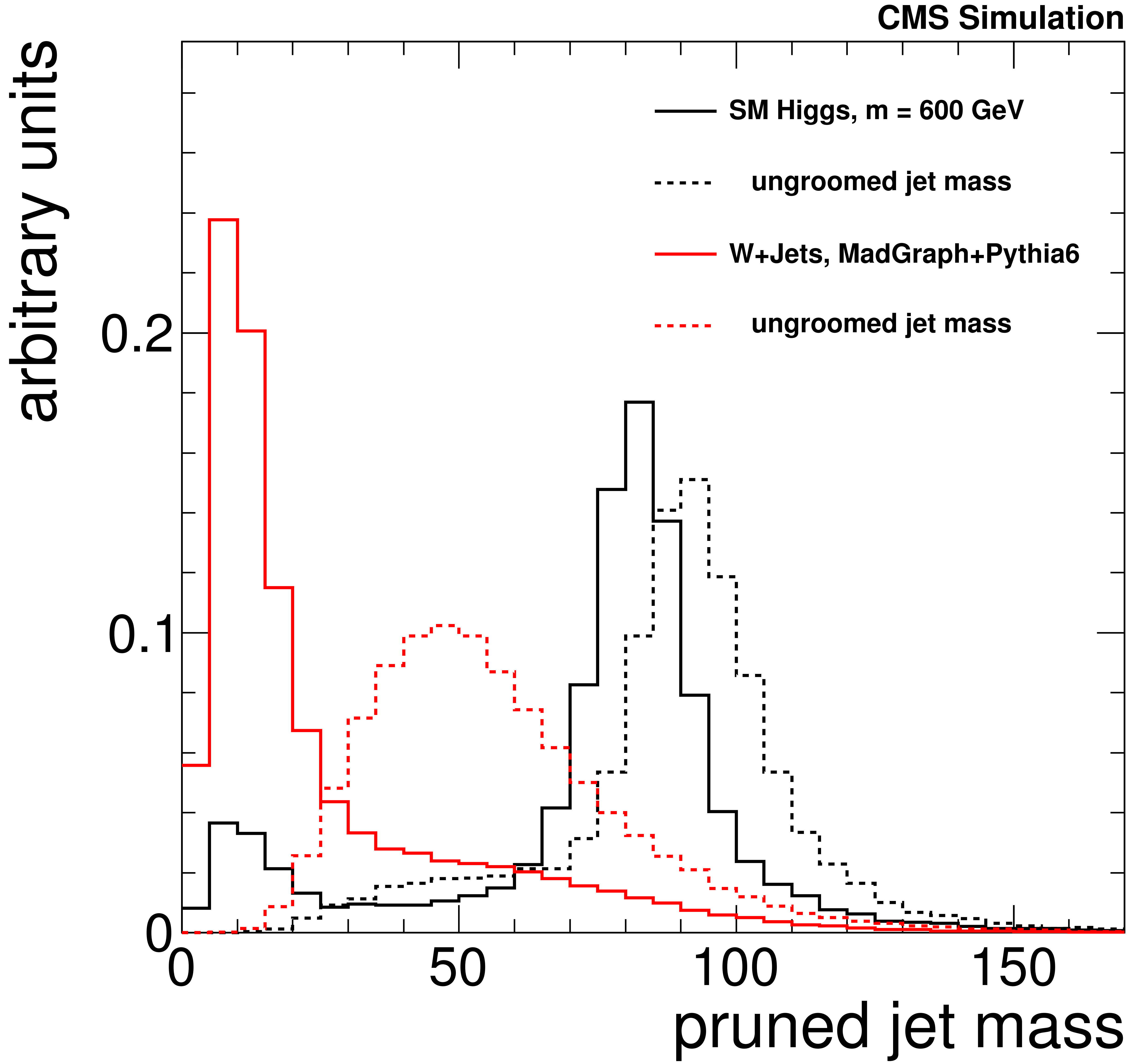
N3LO for VBF have 5-10% effect in our phase space



https://arxiv.org/abs/1506.02660

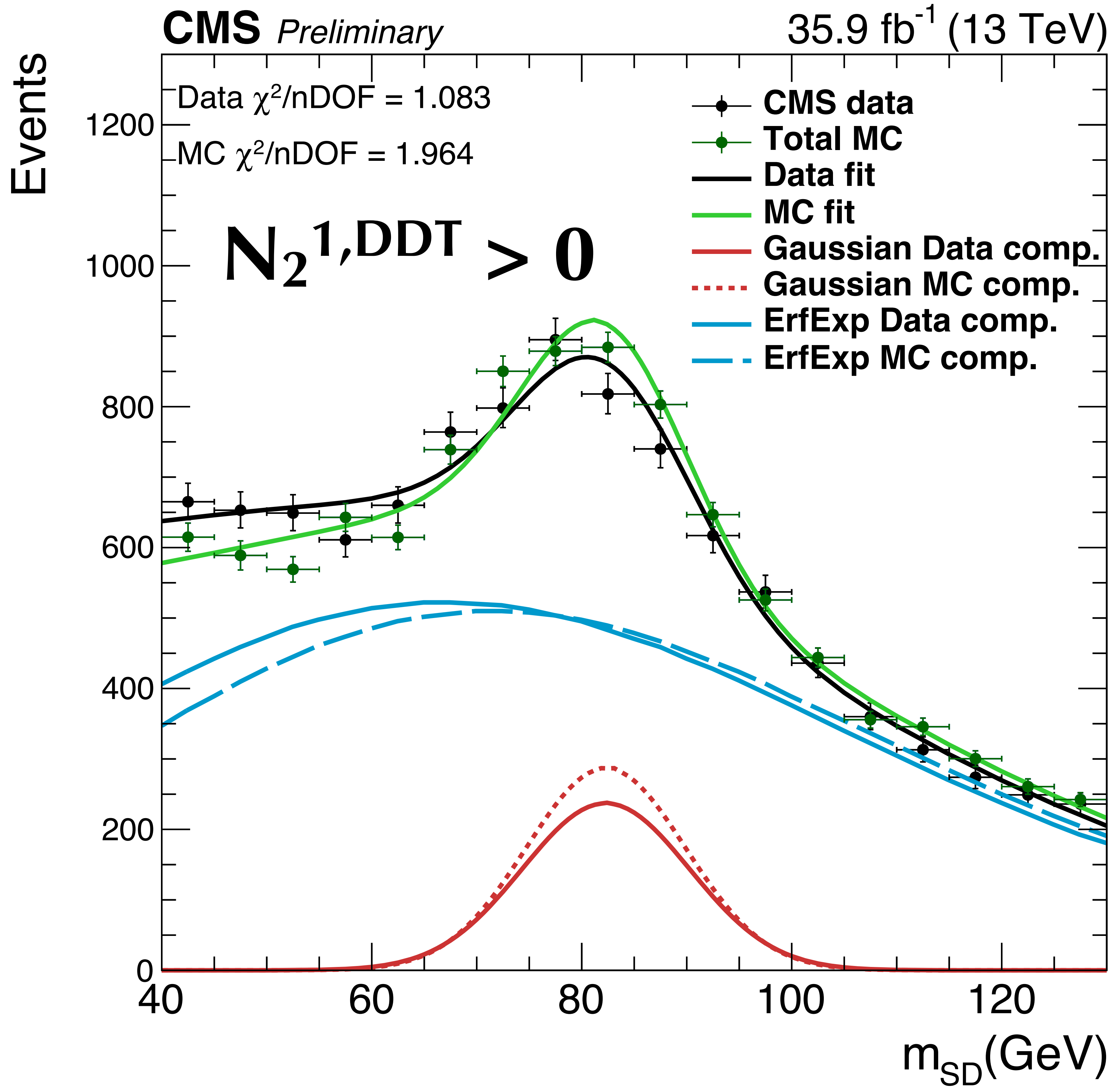
Dreyer, Salam, Zanderighi,Karlberg private communication

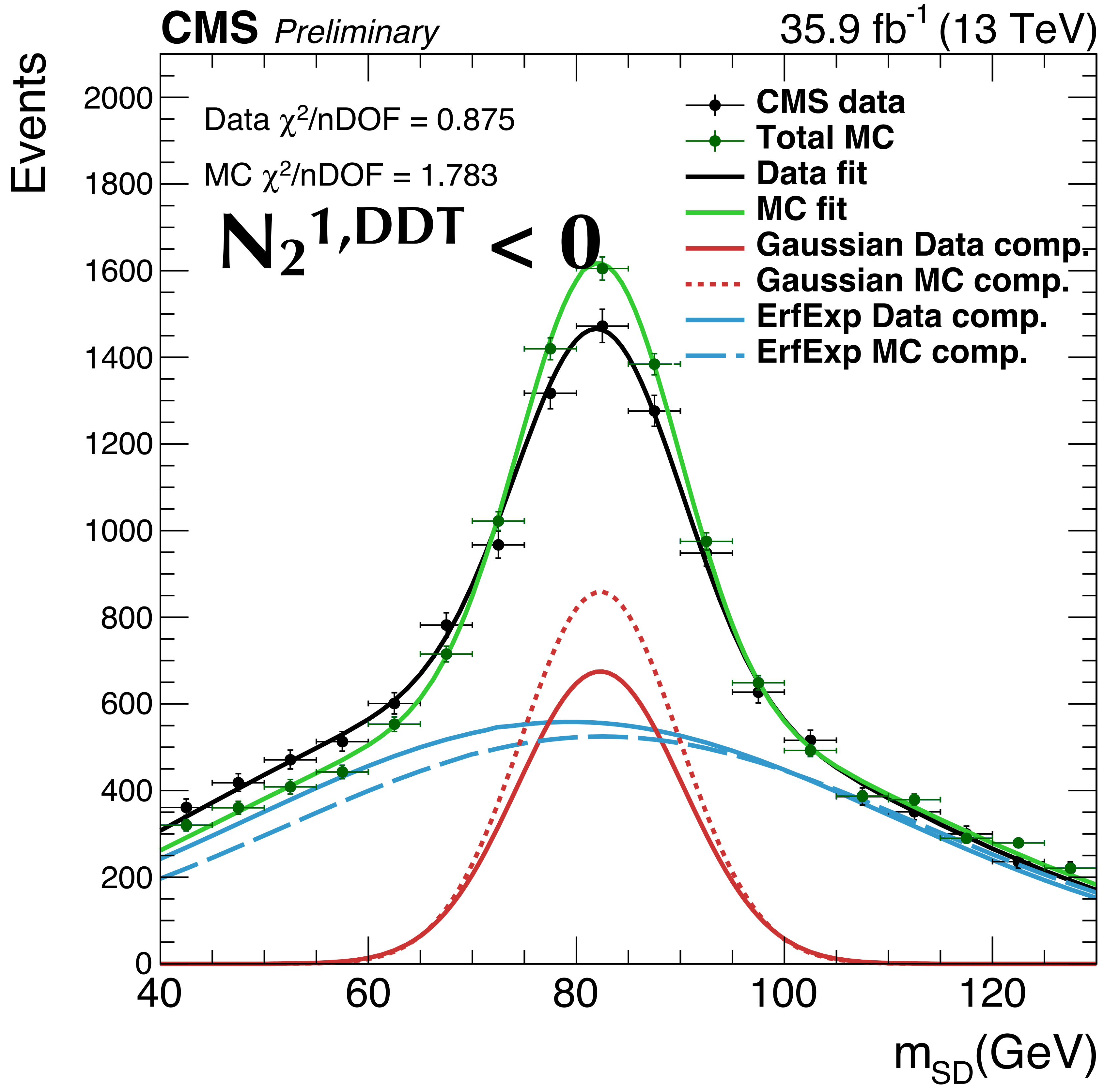
Jet mass grooming



$N_2^{1,DDT}$ efficiency measurement

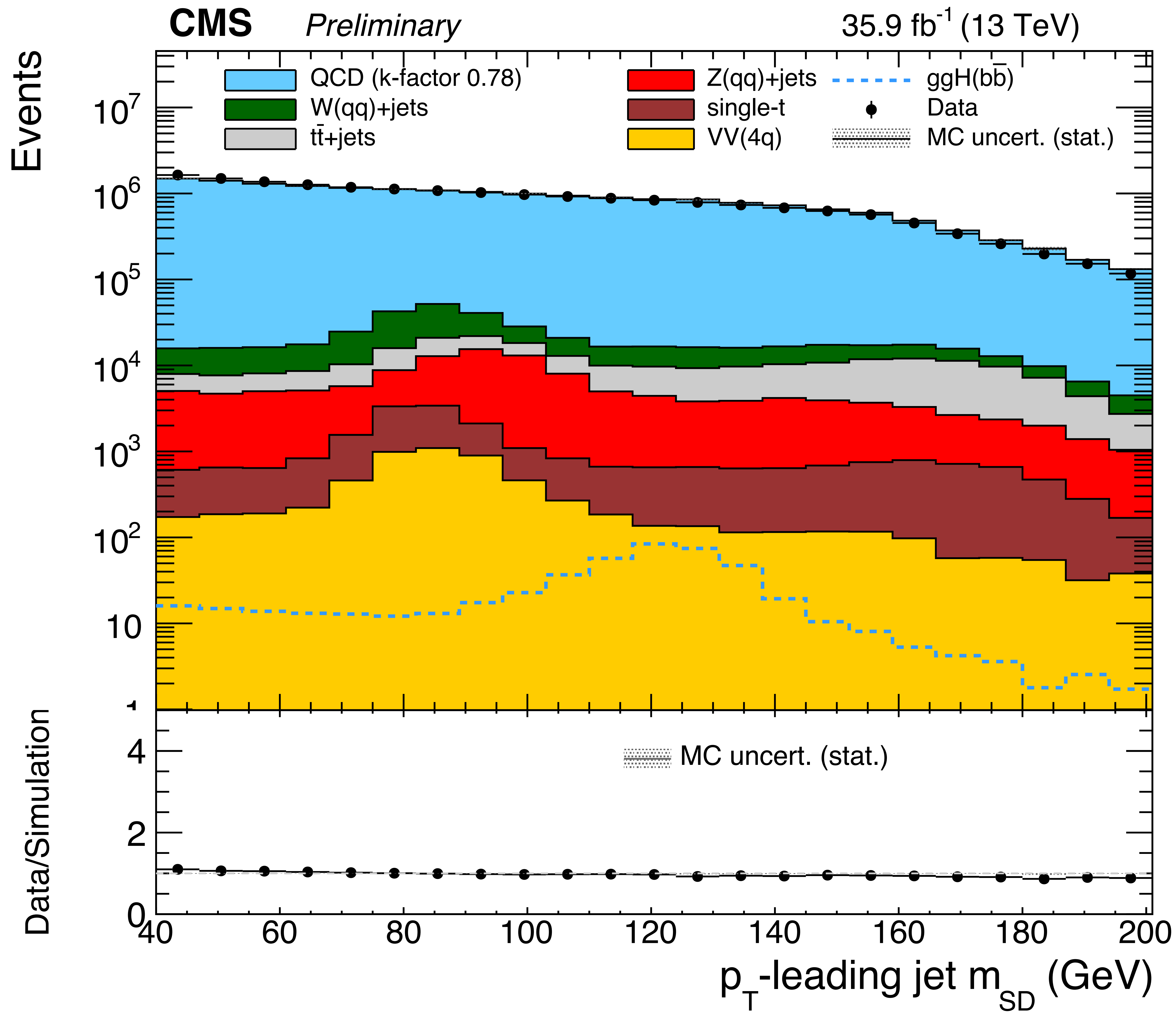
- Efficiency of the $N_2^{1,DDT}$ is measured in data using merged W jets from tt events





- efficiency SF = $\varepsilon_{Data} / \varepsilon_{MC}$ = 0.993 +/- 0.043
- mass scale SF = m_{Data} / m_{MC} = 1.001 +/- 0.004
- mass resolution SF = $\sigma_{Data} / \sigma_{MC}$ = 1.08 +/- 0.09

Background composition



Backgrounds

- Multijets
- tt+jets, normalization from a dedicated control region

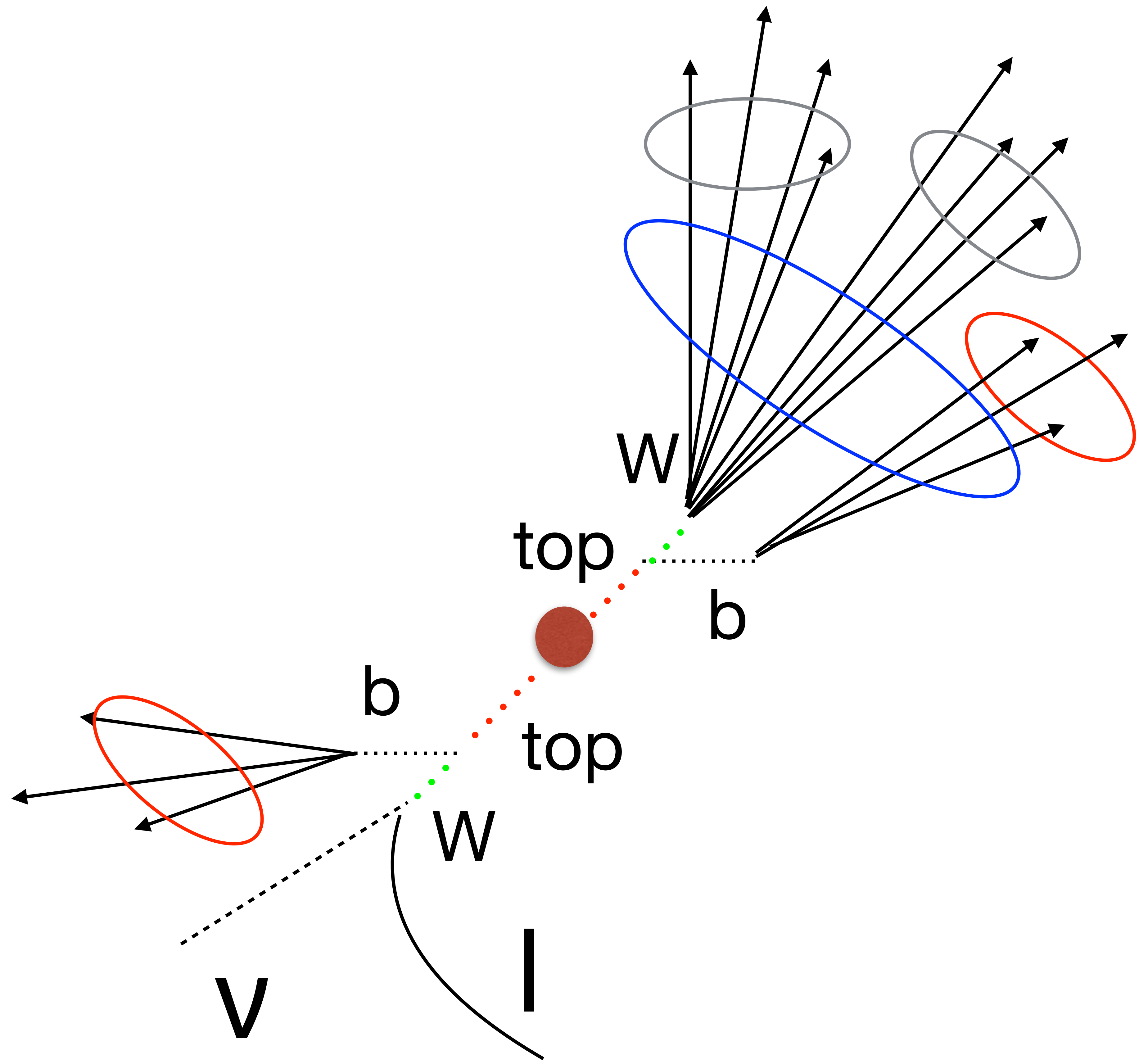
} **data**

- W/Z+jets
- single-t, Diboson

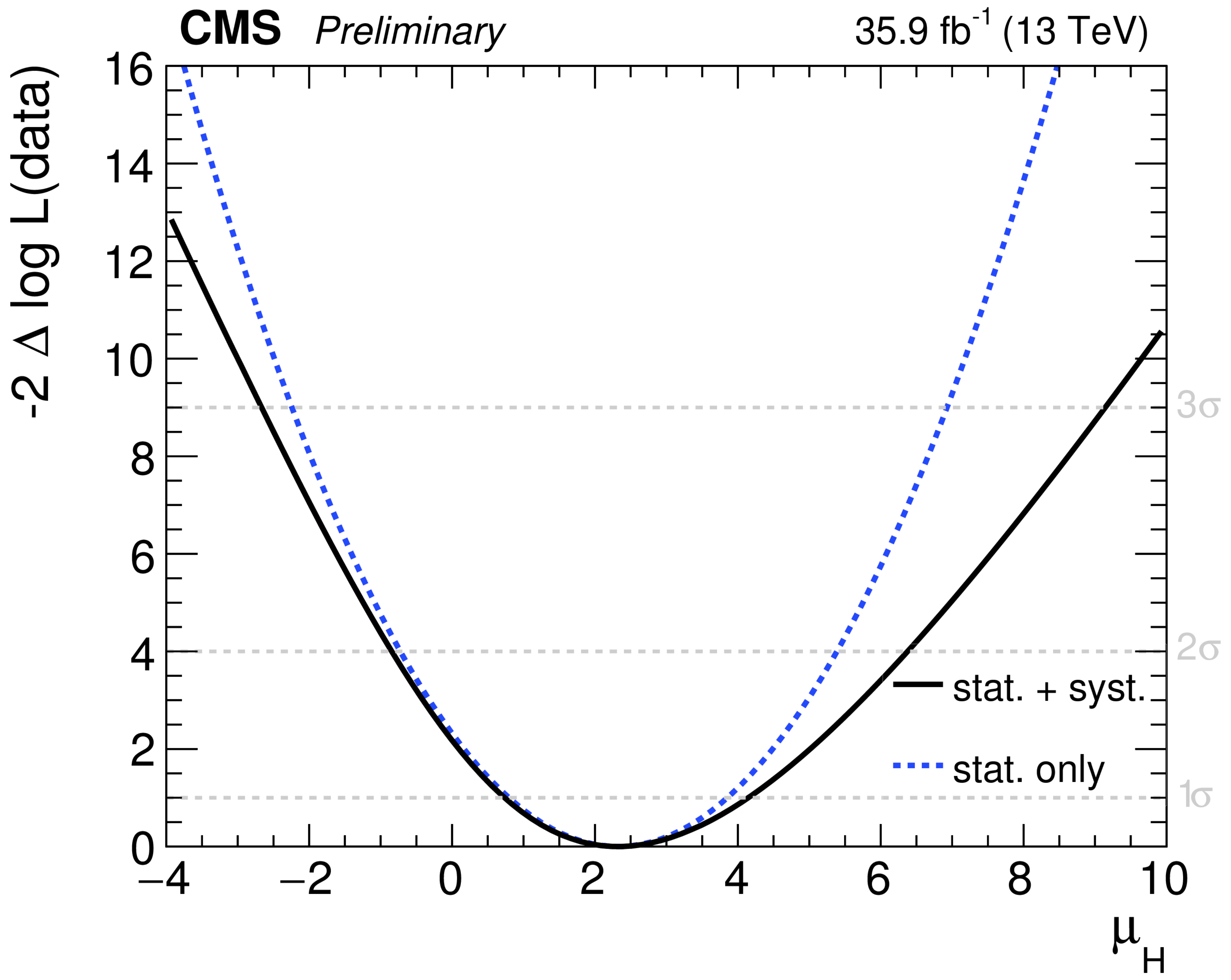
} **MC**

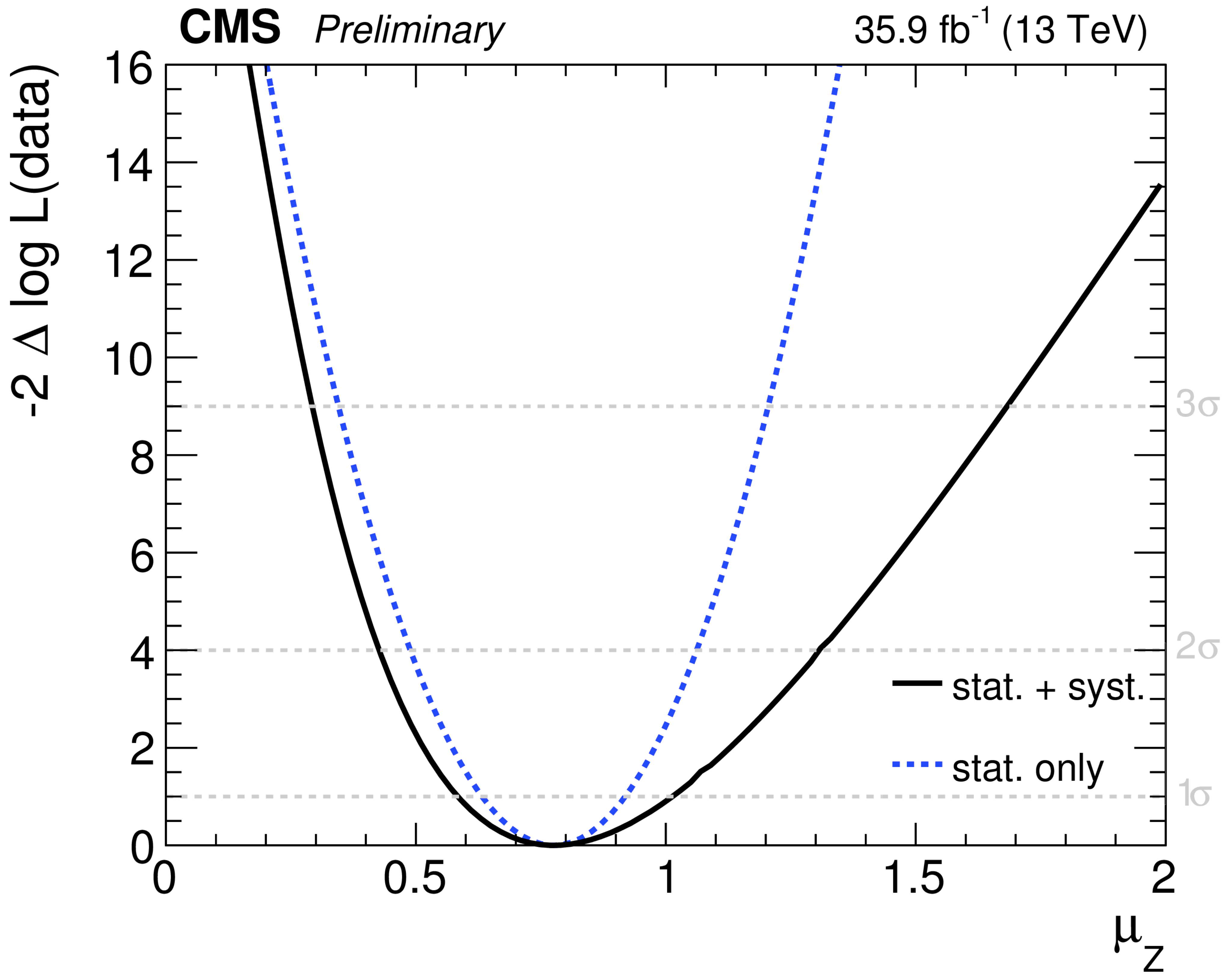
tt enriched control region

- Loose muon with $p_T > 55$ GeV, $|\eta| < 2.1$ in opposite hemisphere:
  - $|\varphi(\mu) - \varphi(\text{AK8 jet})| > 2\pi/3$
- One AK4 PUPPI jet with medium CSVv2 b-tag, with $p_T > 50$ GeV, $|\eta| < 2.5$, and
  - $\Delta R(\text{AK4 b-tag, AK8 jet}) > 0.8$
- Lepton: Veto the presence of identified loose electrons and loose hadronic taus
- Two-prong AK8 PUPPI jet ($p_T > 400$ GeV, $m_{SD} > 40$ GeV and $|\eta| < 2.4$)
  - $N_2^{1,DDT} < 0$



1D LL Scans

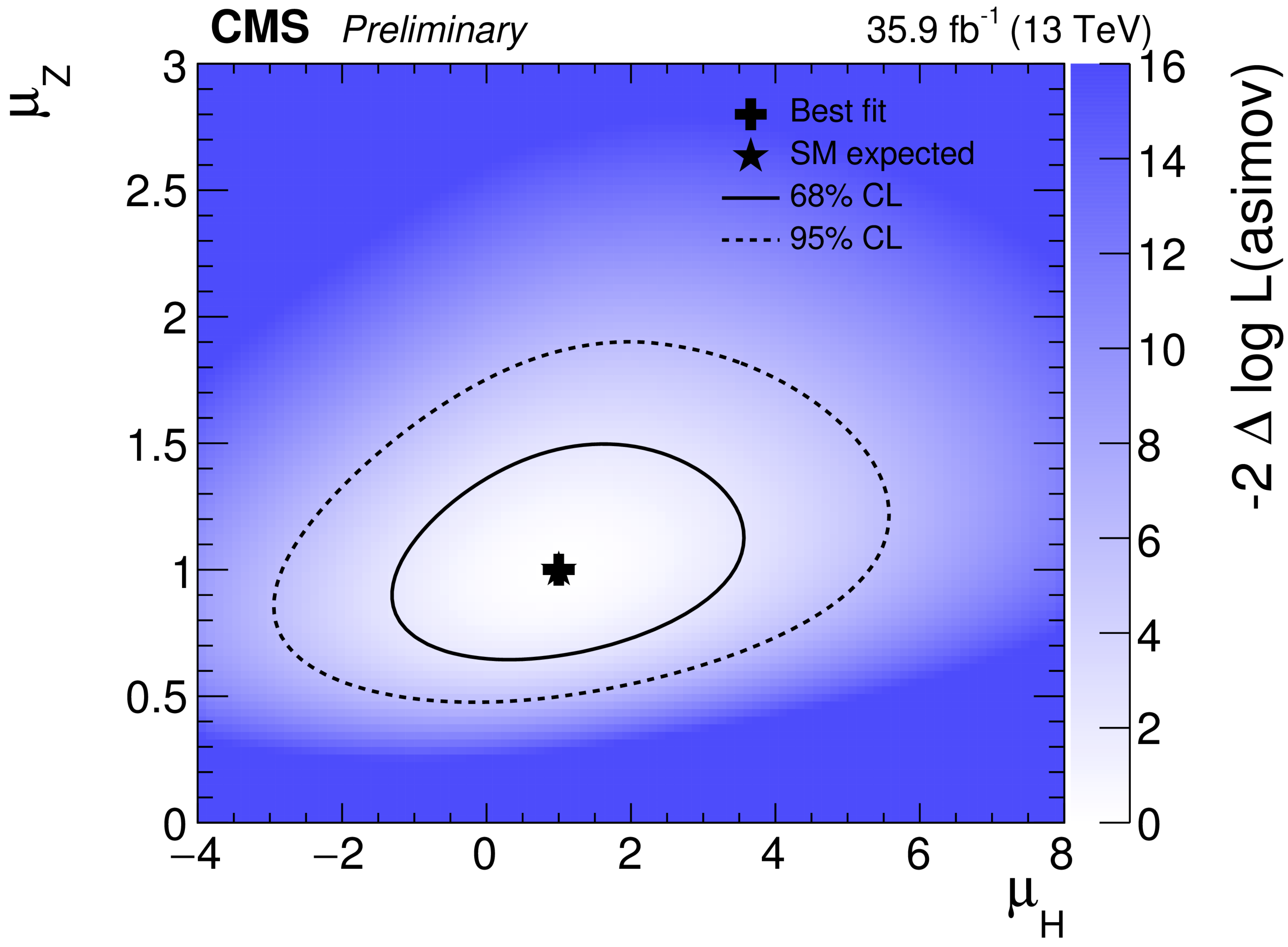




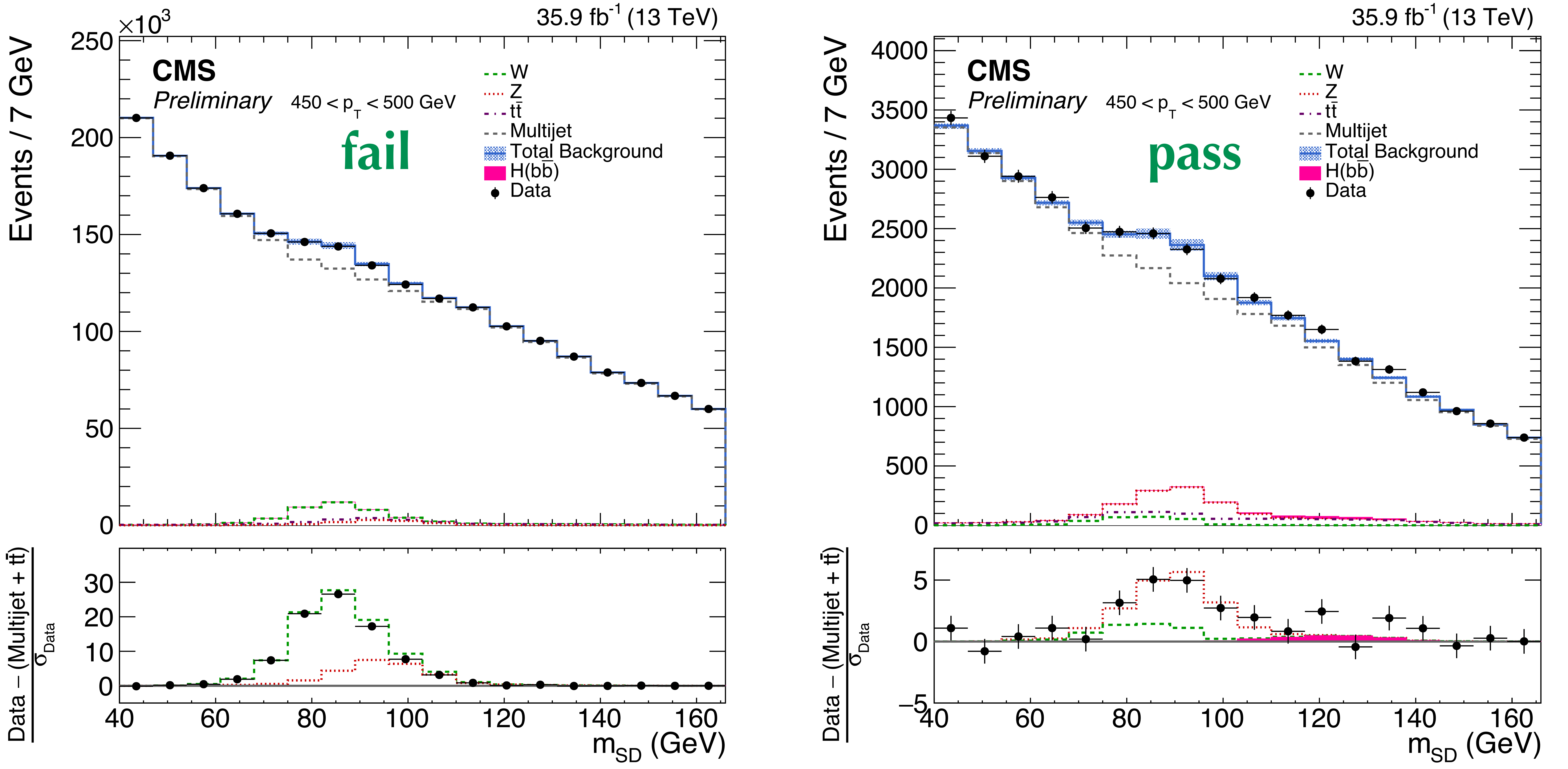
μ_Z = 0.78 −0.14/+0.14 (stat.) −0.13/+0.19 (syst.)

μ = 2.3 −1.5/+1.5 (stat.) −0.4/+1.0 (syst.)

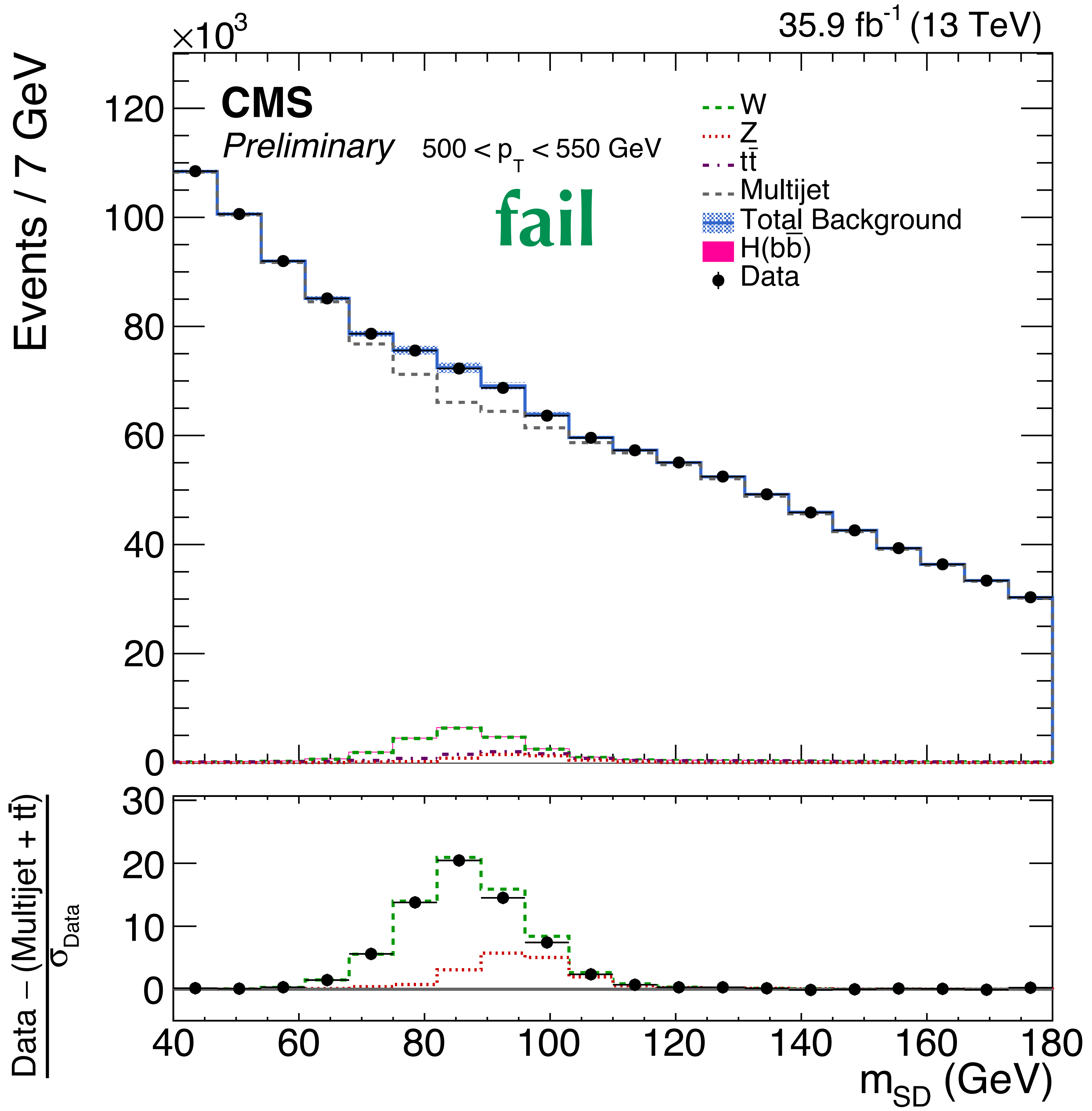
2D LL Scan (Asimov)

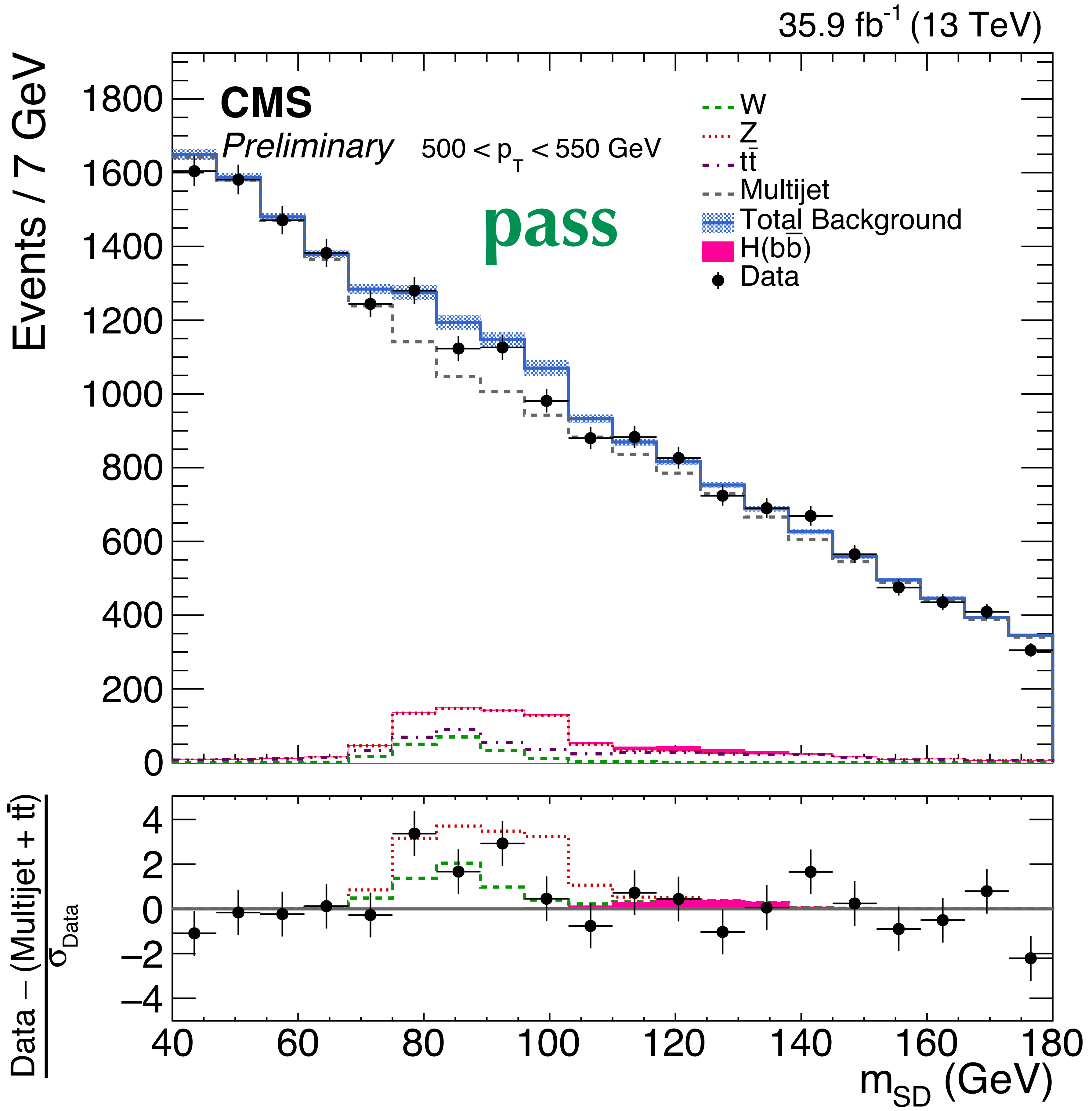


Results

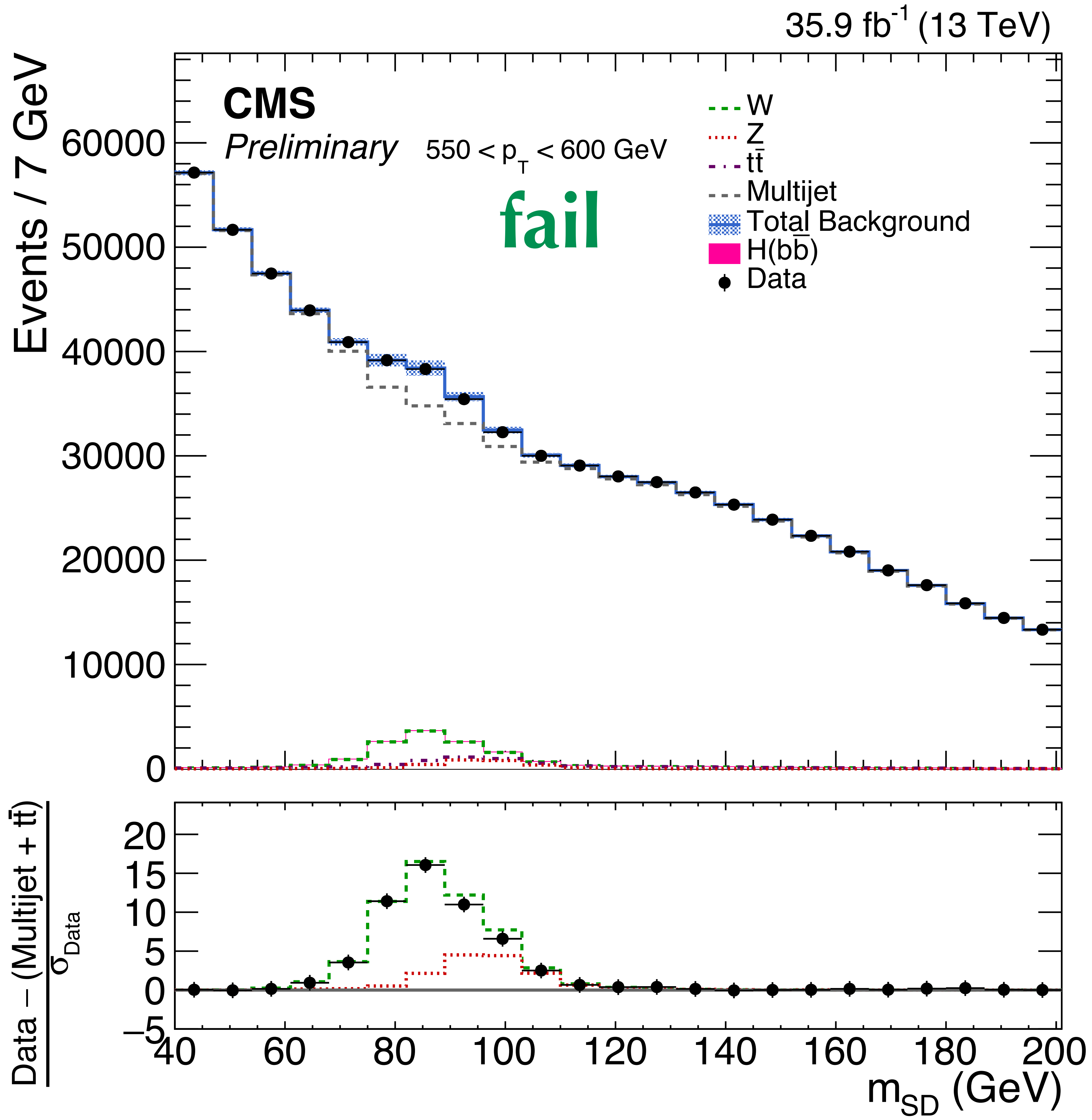

35.9 fb^{-1} (13 TeV)
CMS
Preliminary
450 < p$_T$ < 500 GeV
fail
W
Z
t$\bar{t}$
Multijet
Total Background
H(b$\bar{b}$)
Data
Events / 7 GeV
Data − (Multijet + t$\bar{t}$) / σ$_{Data}$
m$_{SD}$ (GeV)
35.9 fb^{-1} (13 TeV)
CMS
Preliminary
450 < p$_T$ < 500 GeV
pass
W
Z
t$\bar{t}$
Multijet
Total Background
H(b$\bar{b}$)
Data
Events / 7 GeV
Data − (Multijet + t$\bar{t}$) / σ$_{Data}$
m$_{SD}$ (GeV)

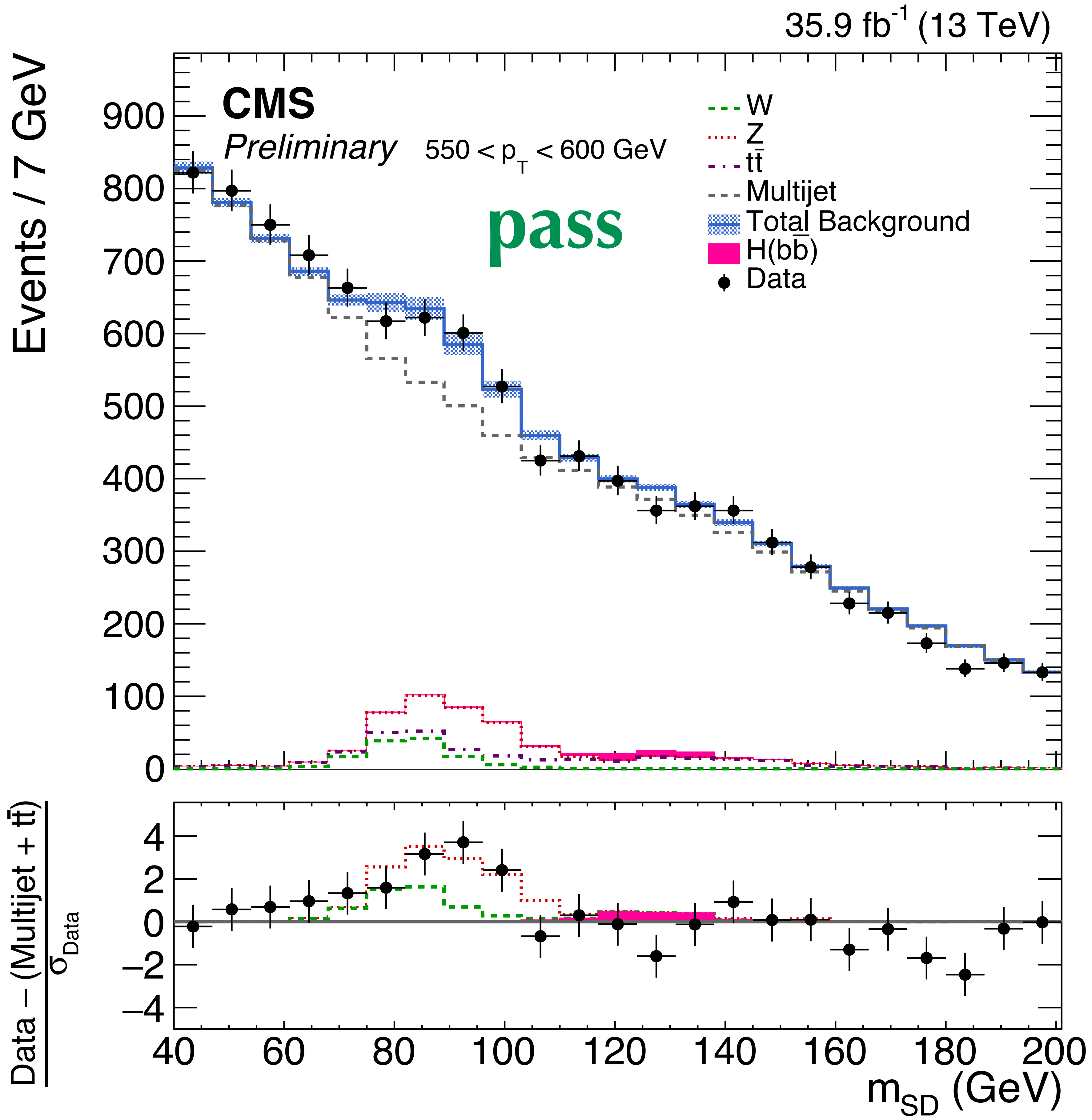
Signal+background fit



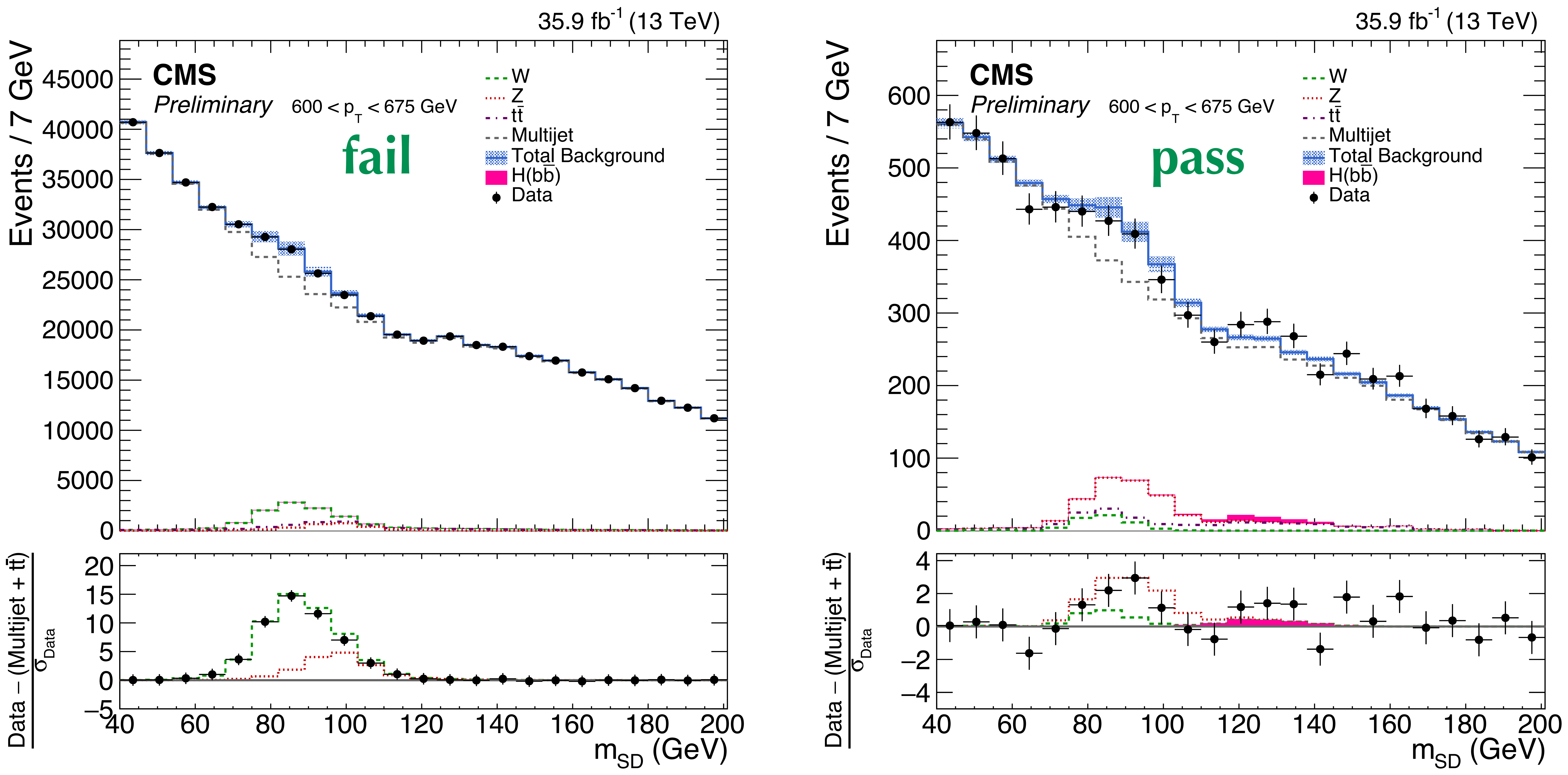


Signal+background fit

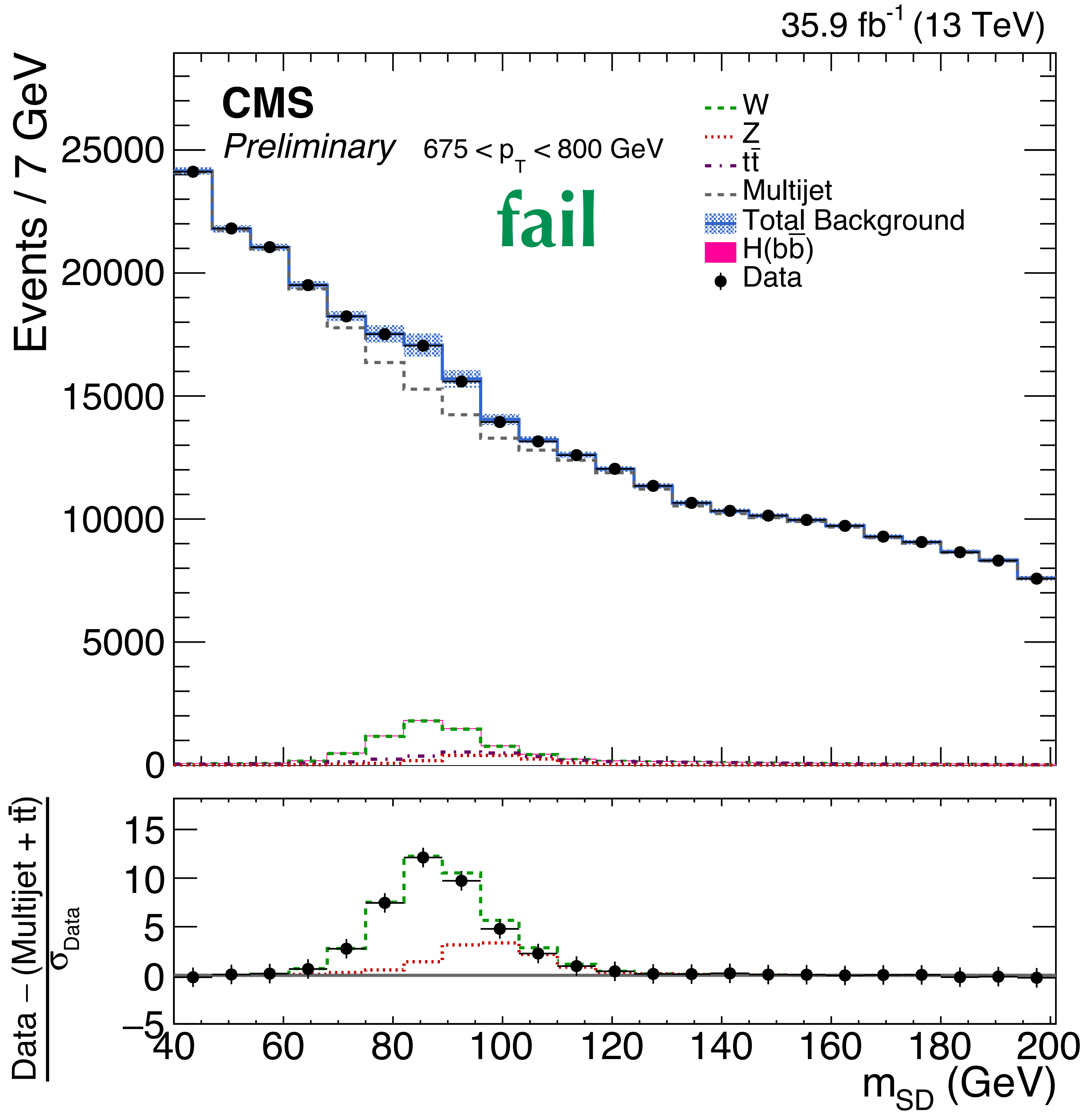


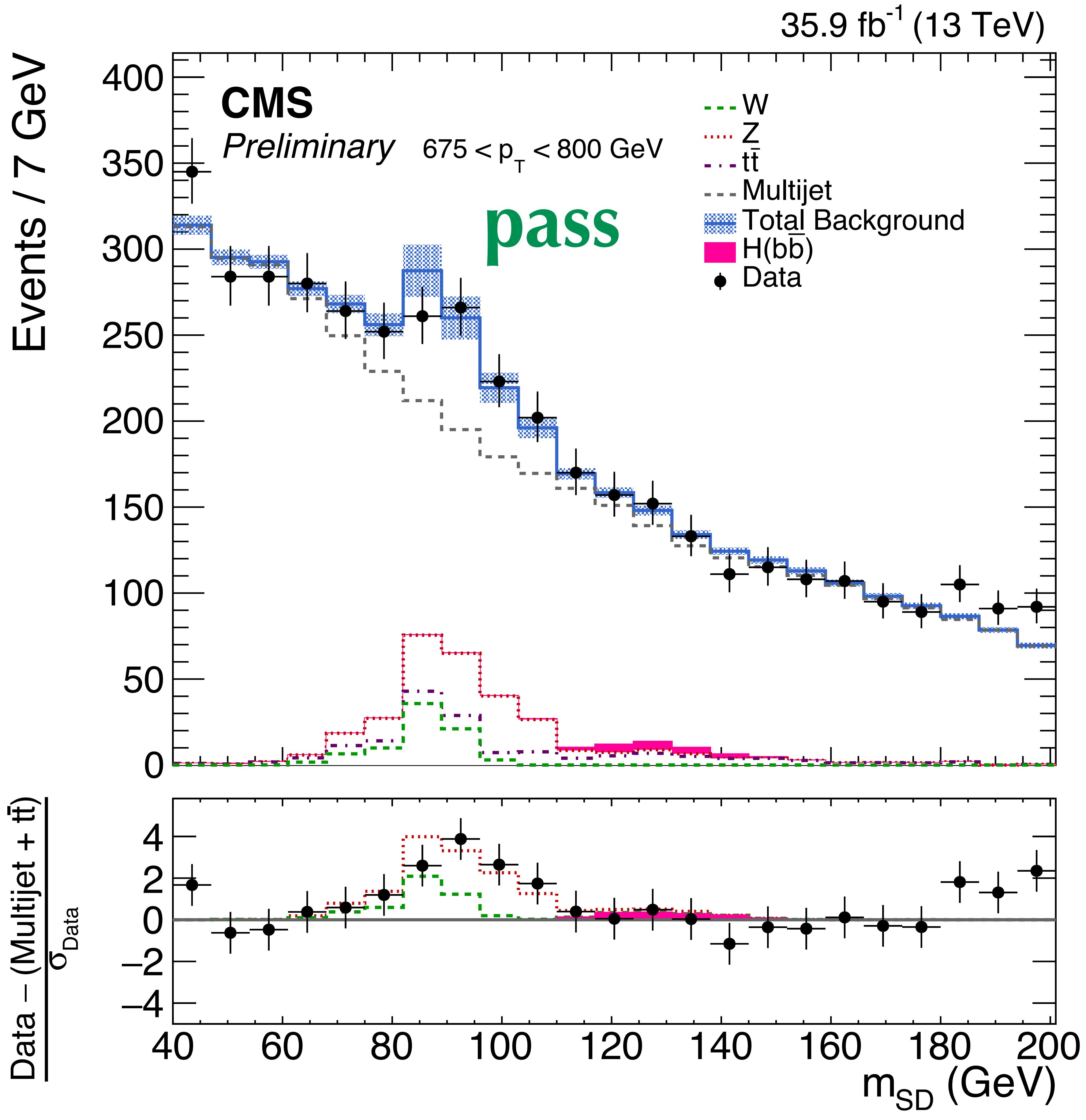


Signal+background fit

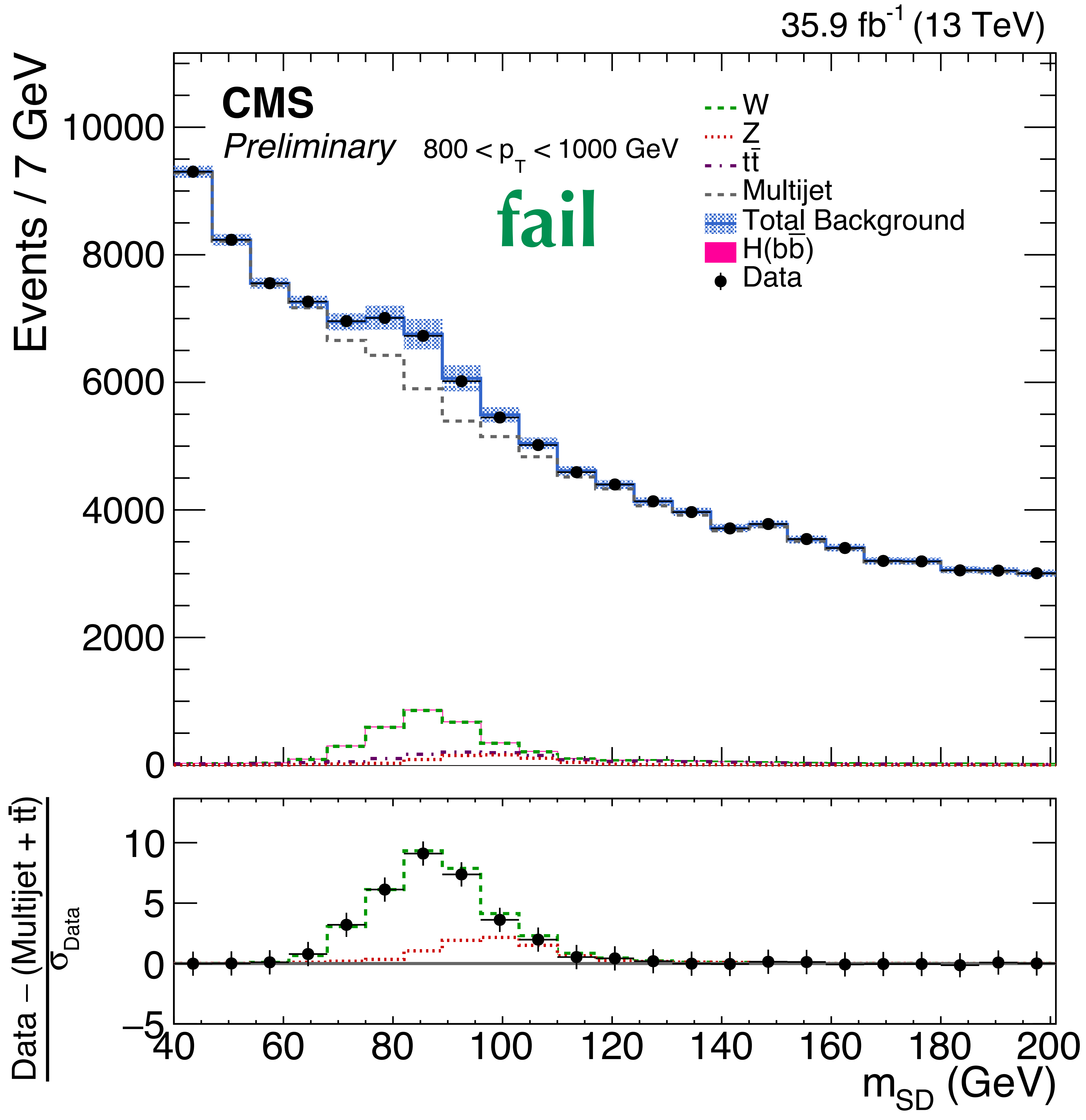


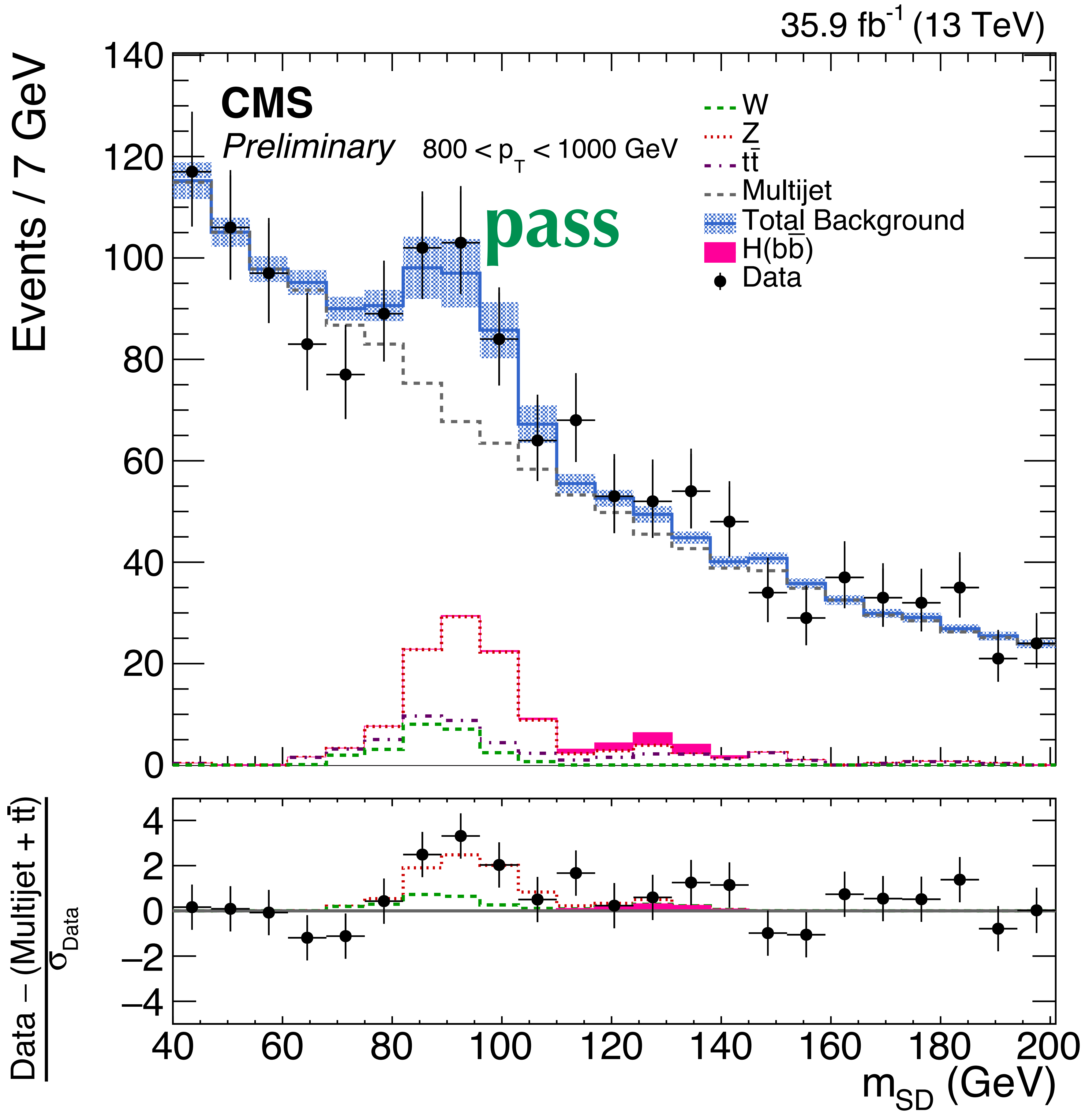
Signal+background fit





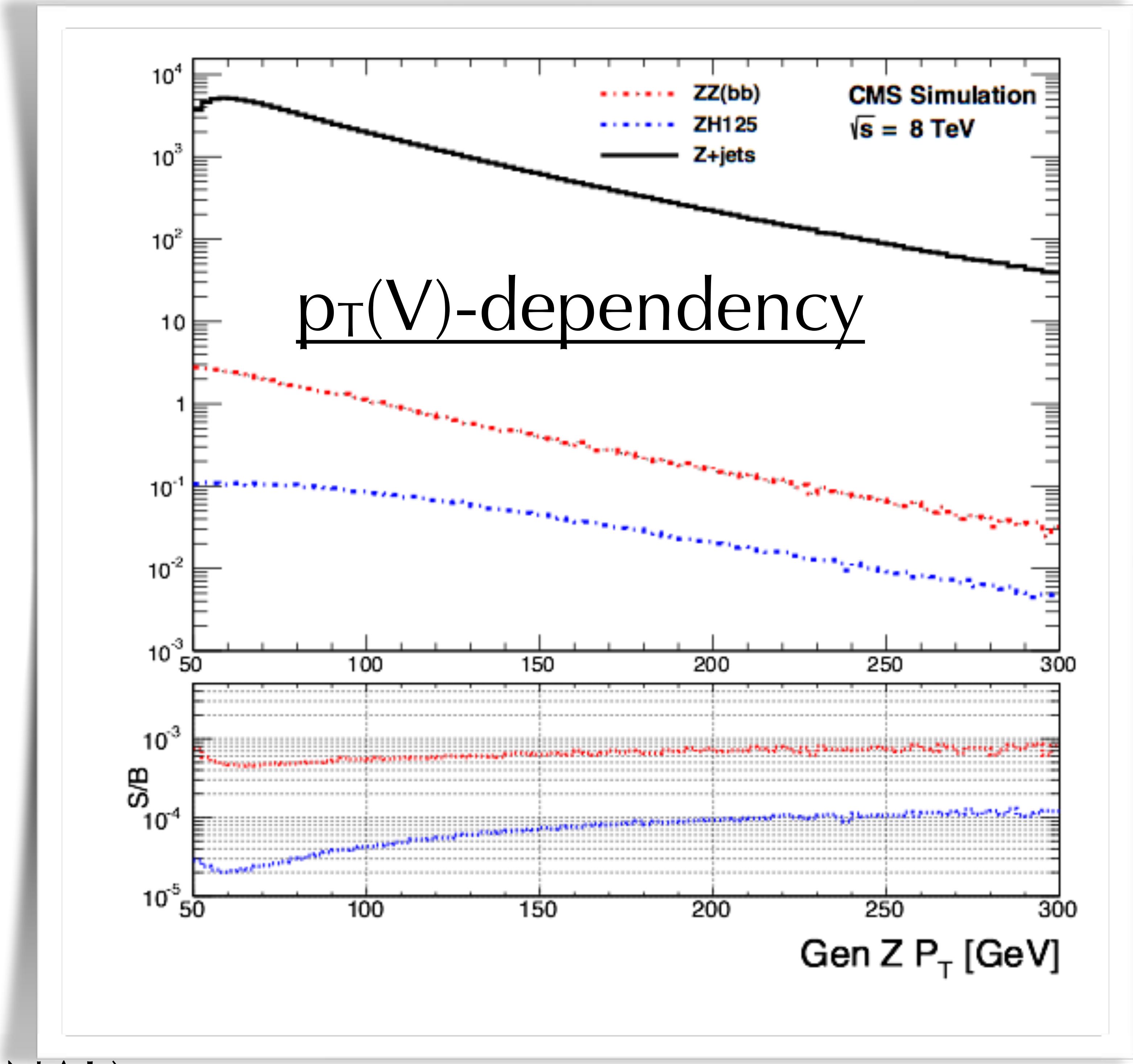
Signal+background fit

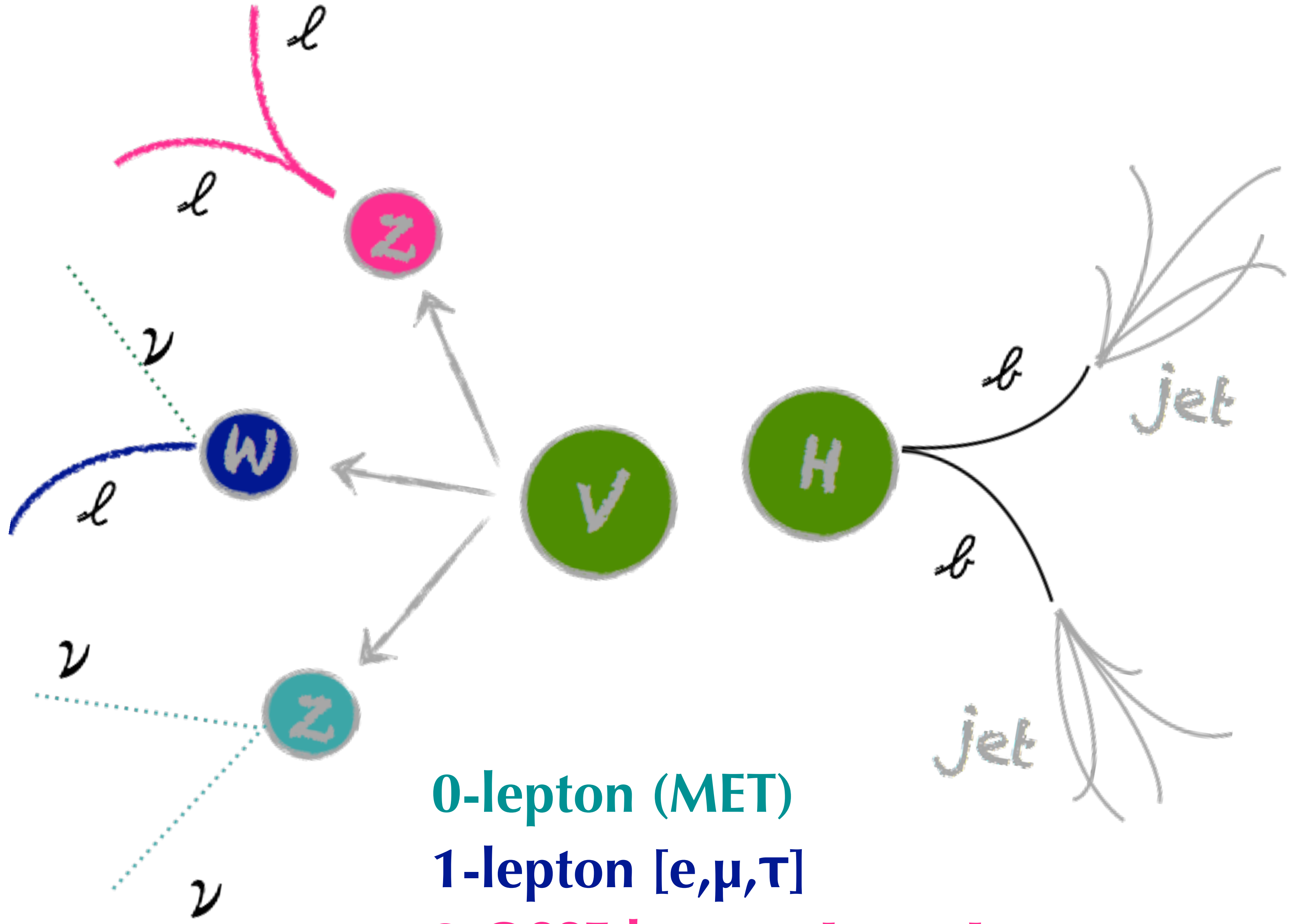




VH, Event Topology

- $H \to b\bar{b}$ at LHC is searched in events where H is produced in association with a W or Z boson with high boost (~ 100 GeV)
  - events are triggered by the leptonic decay of the W/Z (e, μ, MET)
  - multi-jet QCD background is highly suppressed

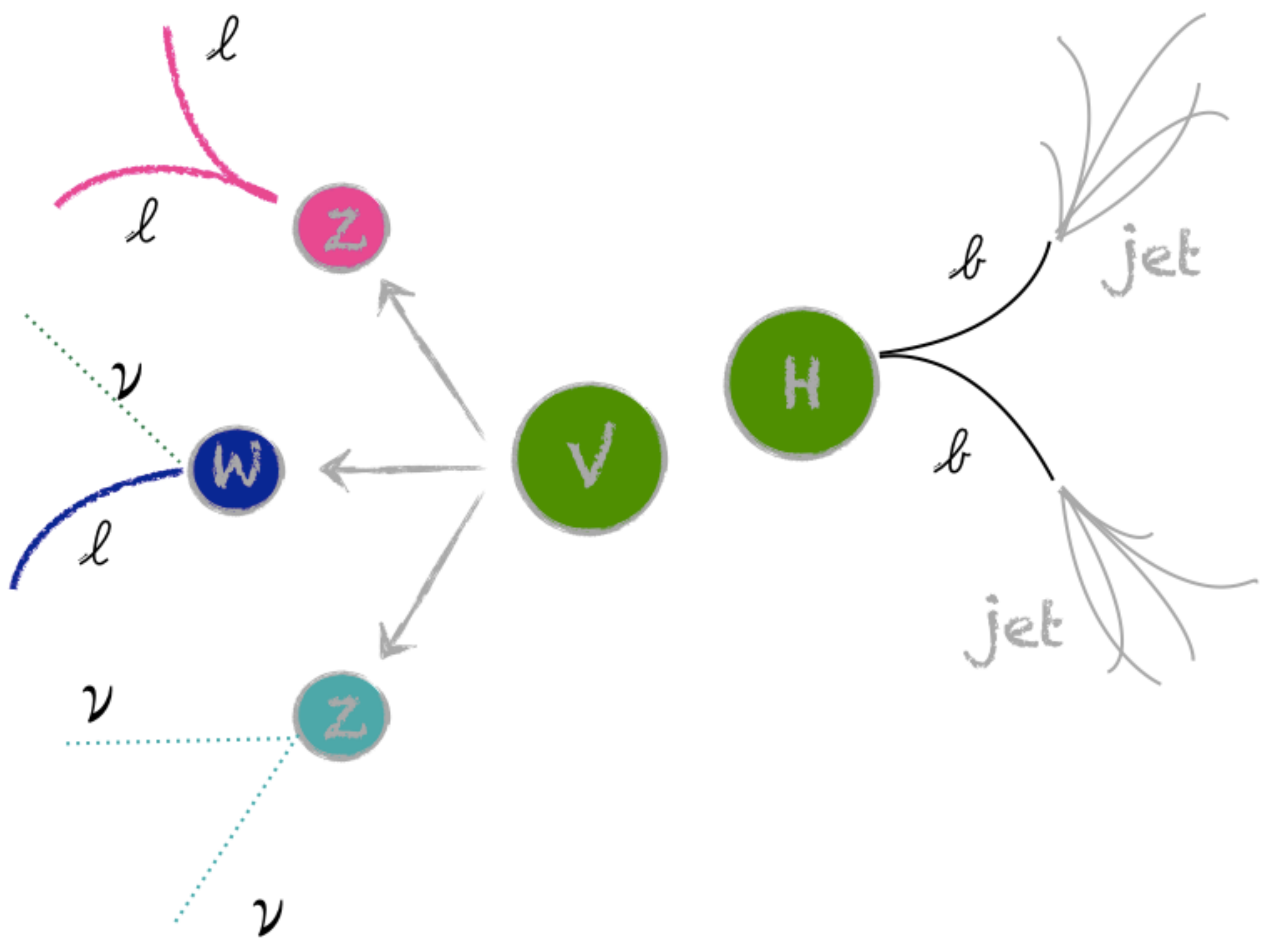




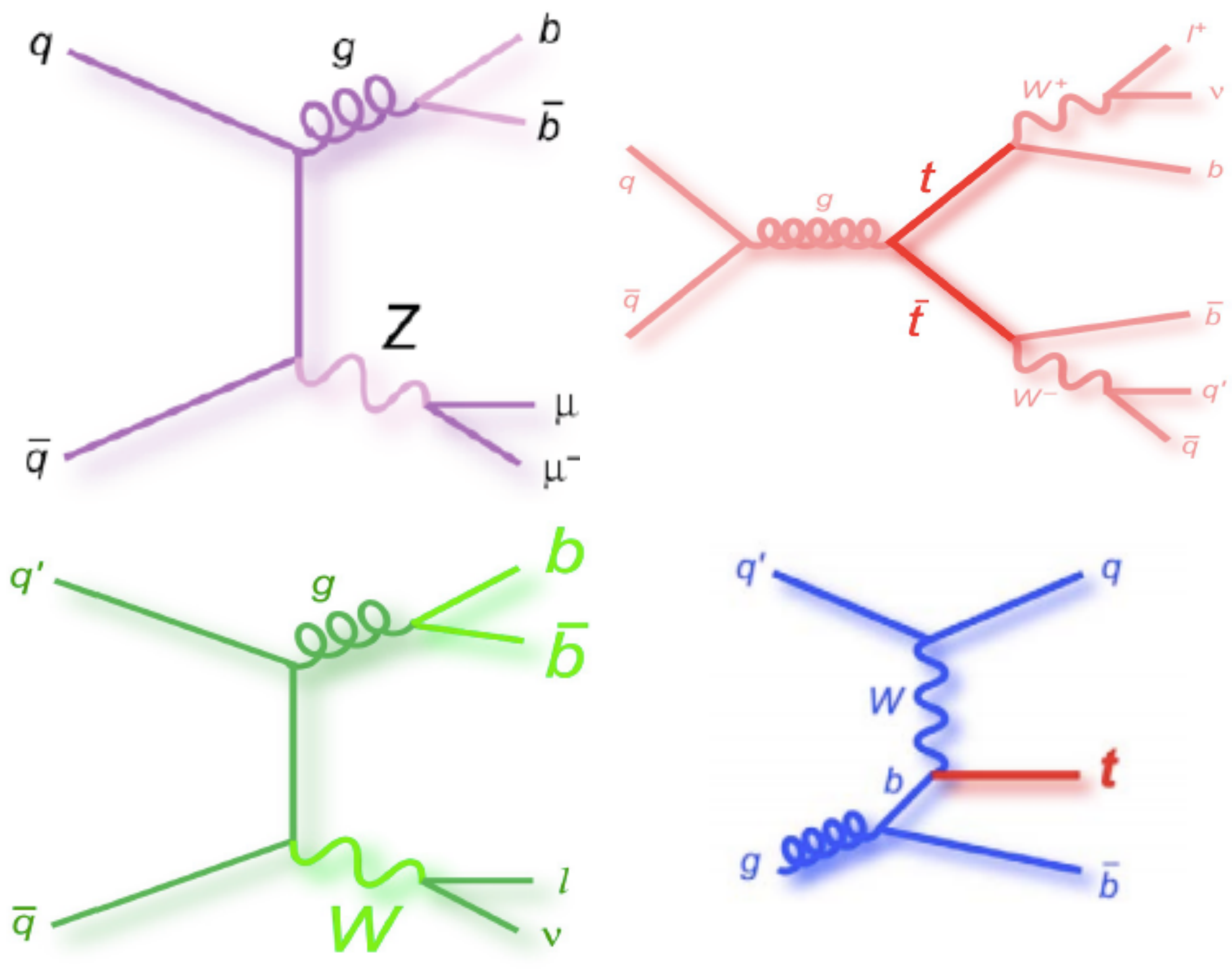
Quick look at the backgrounds

VH example

signal



irreducible backgrounds

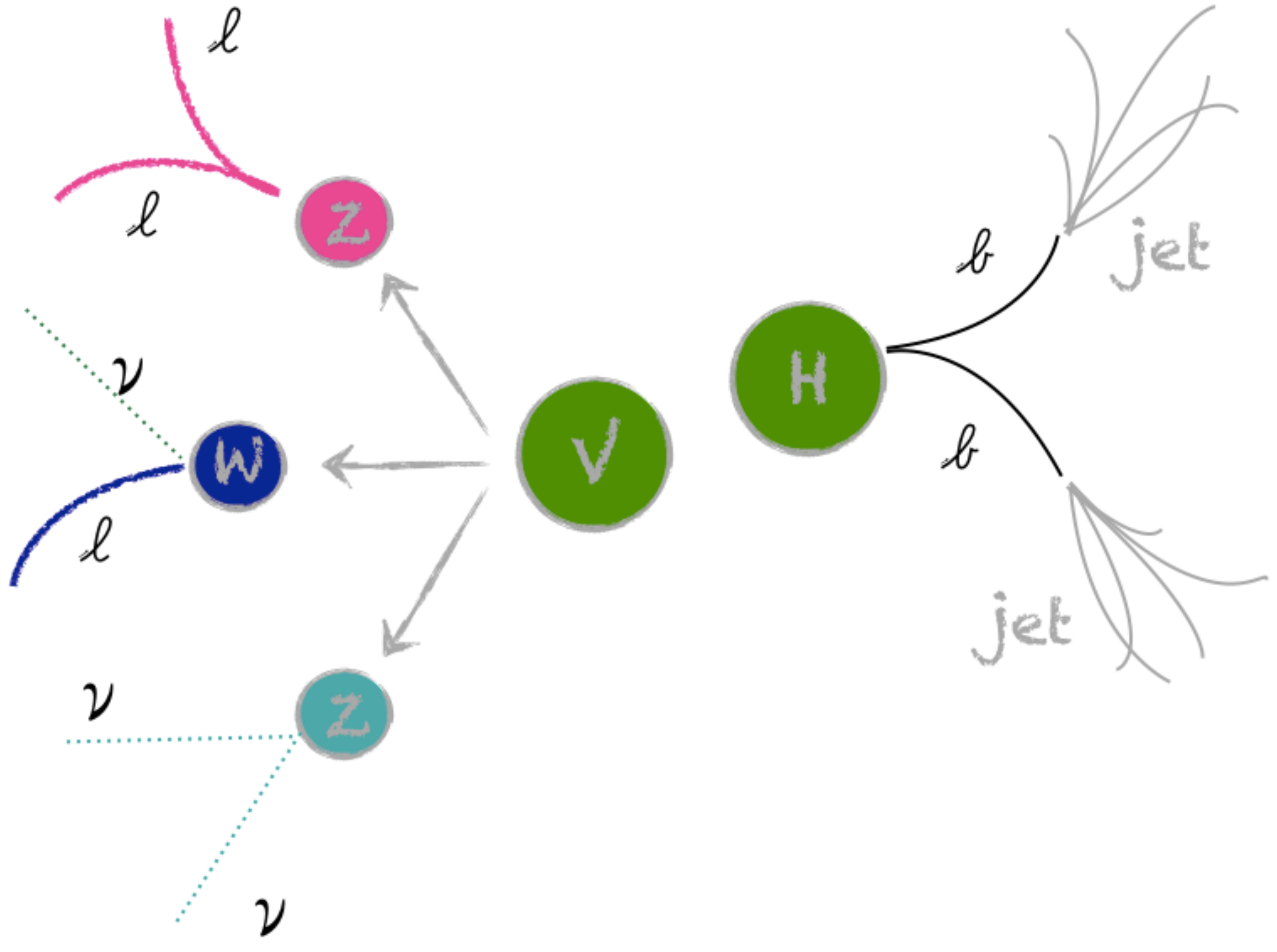


0-lepton (MET)
1-lepton [e,μ,τ]
2-OSSF leptons [ee,μμ]

Quick look at the backgrounds

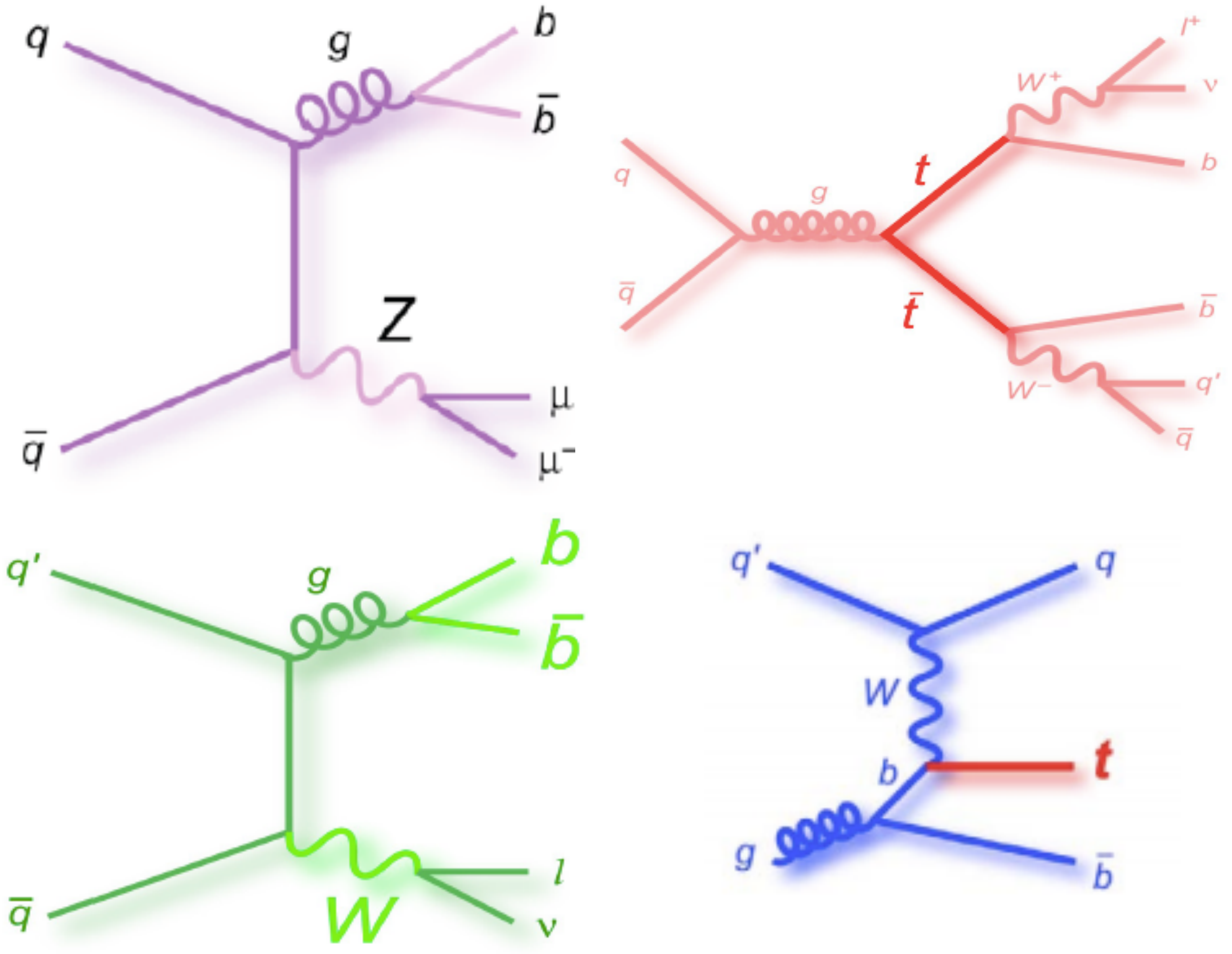
VH example

signal



0-lepton (MET)
1-lepton [e,μ,τ]
2-OSSF leptons [ee,μμ]

irreducible backgrounds

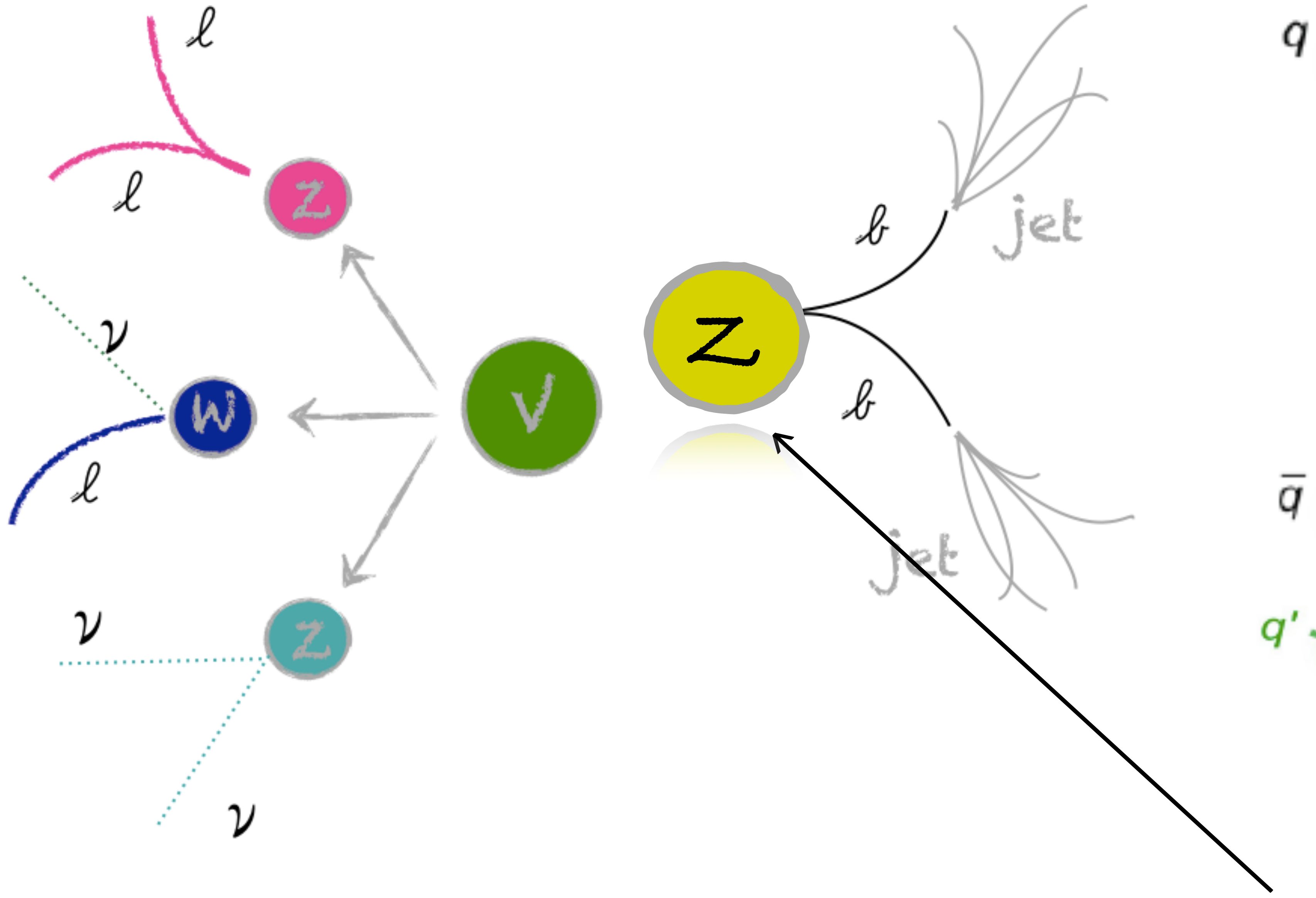


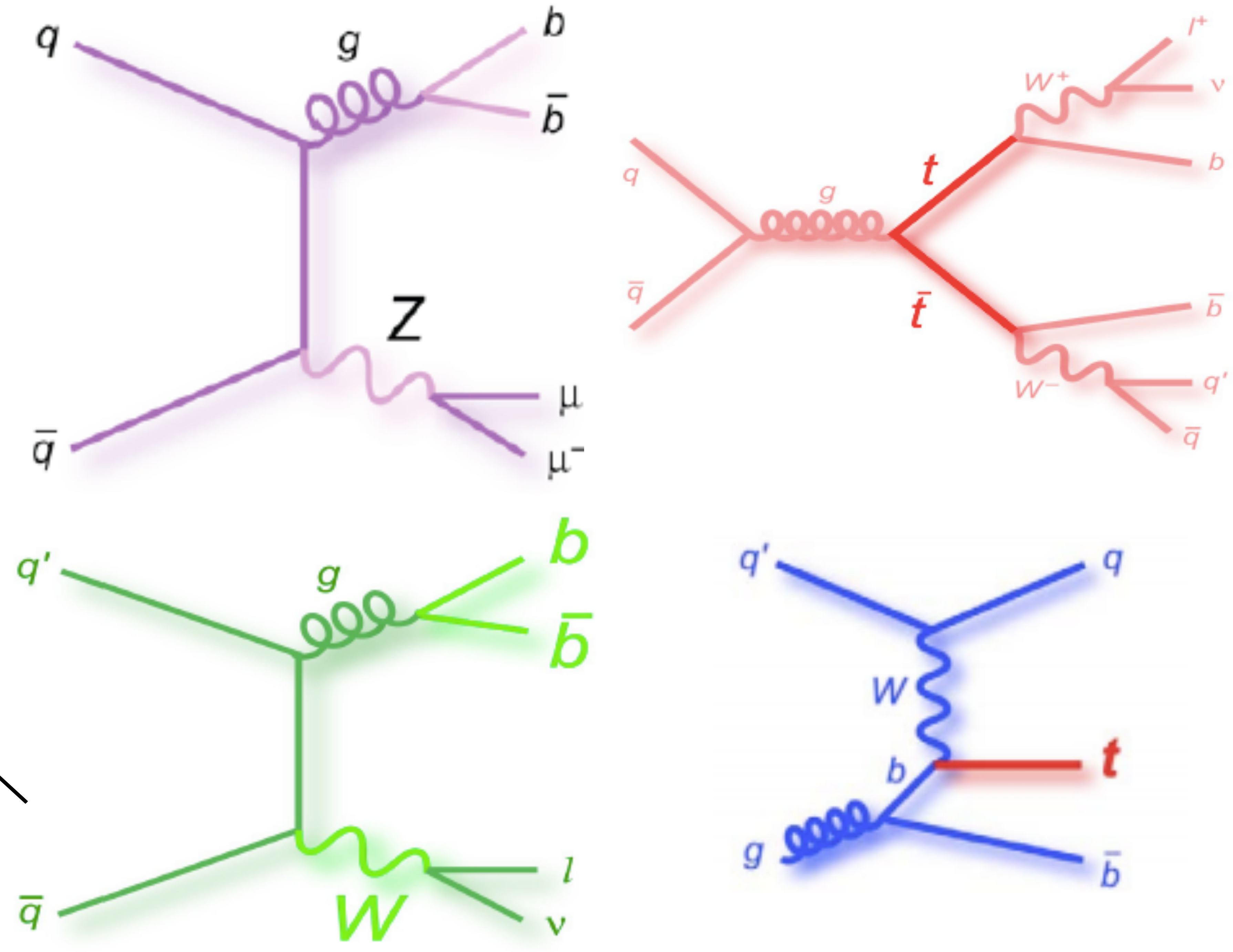
and diboson, of course

Quick look at the backgrounds

VH example

irreducible backgrounds



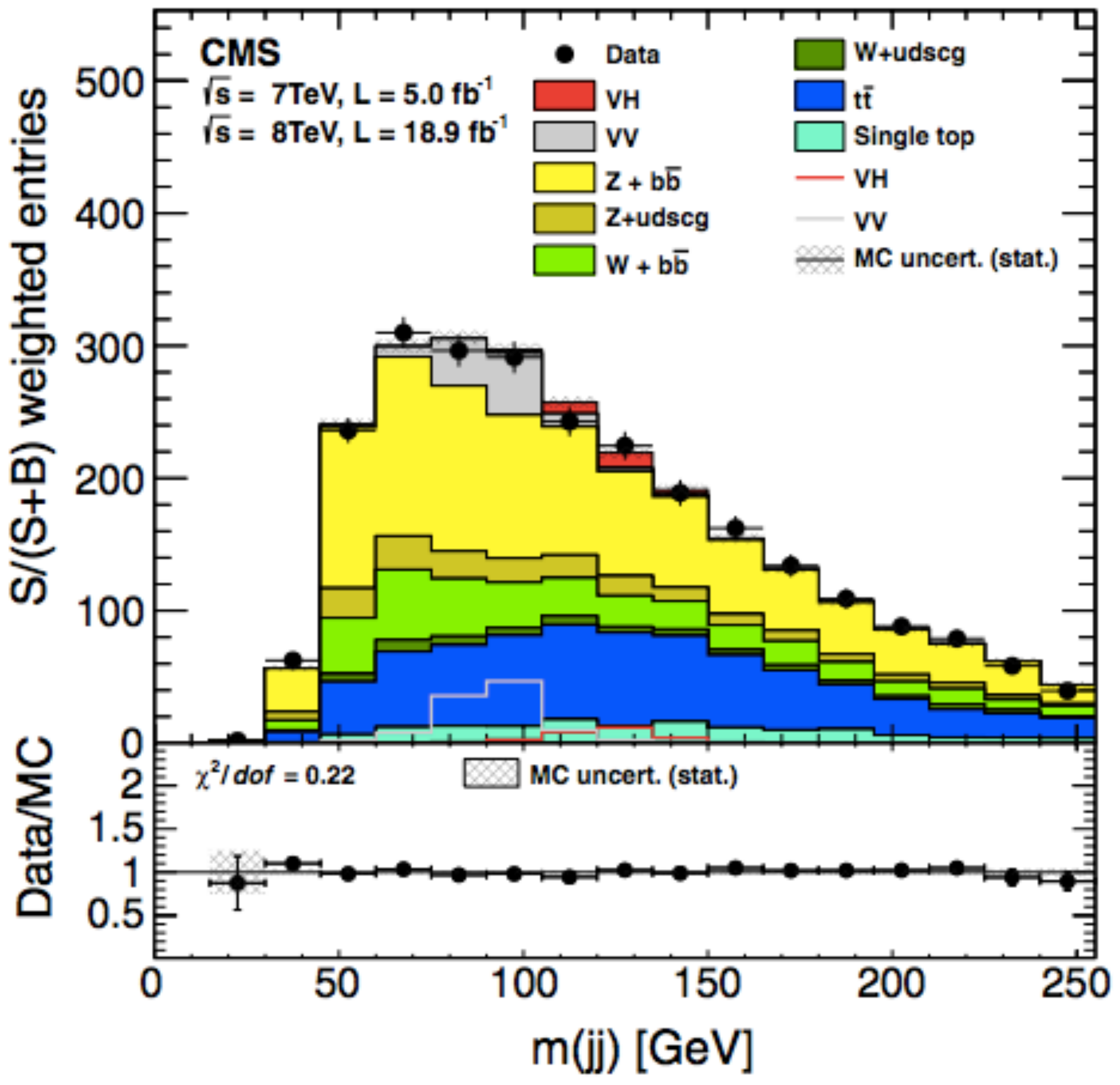


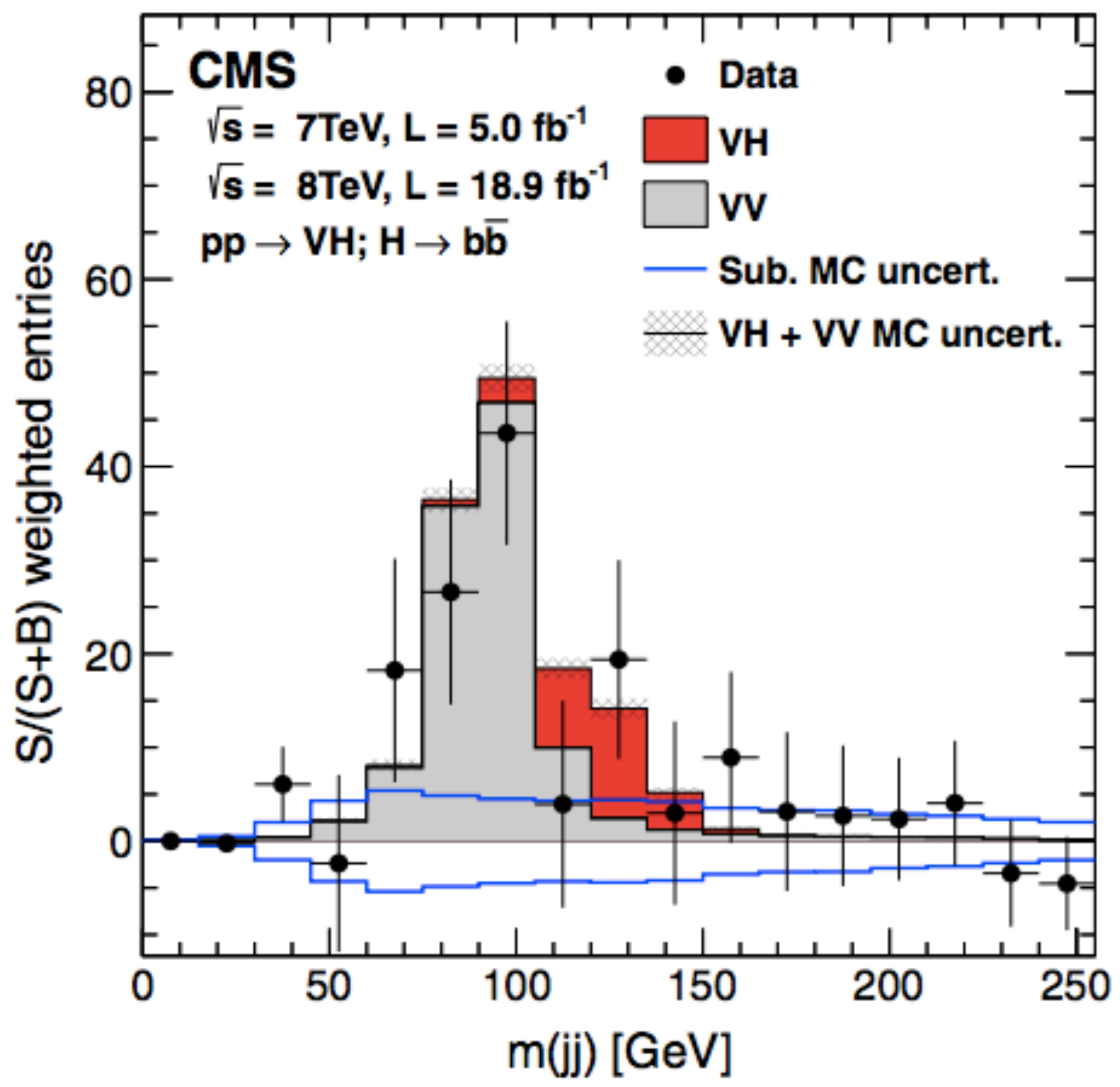
0-lepton (MET)
1-lepton [e,μ,τ]
2-OSSF leptons [ee,μμ]

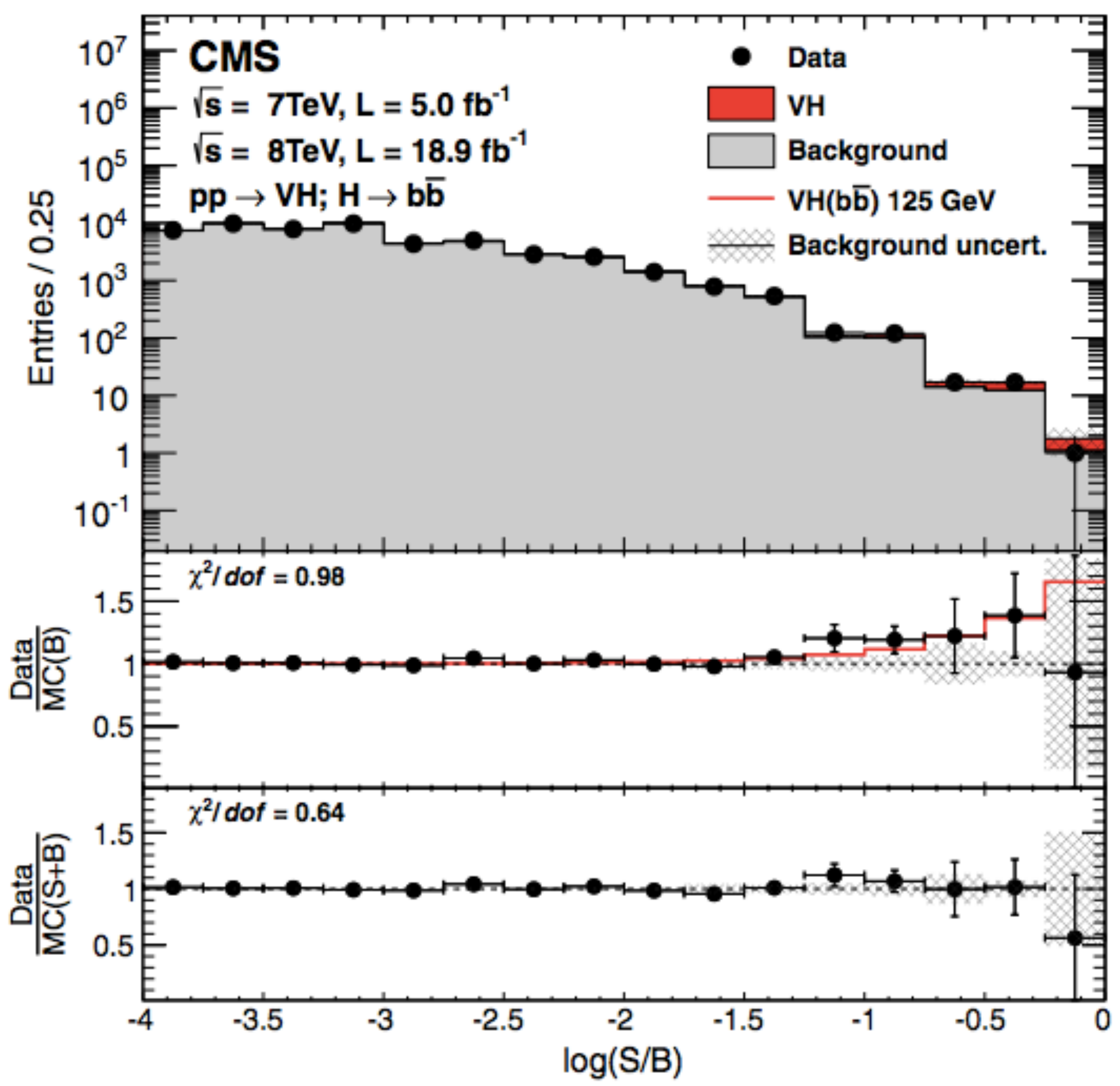
and diboson, of course

Phys. Rev. D 89 (2014)

VH, Analysis Strategy







Key points:

1. Extract normalization for the dominant backgrounds from the data V+0b/1b/2b and top pair production
2. b-jet energy specific corrections (**regression**)
3. A multivariate analysis, BDT

Phys. Rev. D 89 (2014)

VH, Results

VH($b\bar{b}$) reported an excess of 2.1 σ in agreement with SM H expectation at 125 GeV

- ✓ $\mu = \sigma/\sigma_{SM} = 1.0 \pm 0.5$
- ✓ All modes are compatible
- ✓ most sensitive result, to be compared to

CDF $\mu = \sigma/\sigma_{SM} = 2.5 \pm 1.0$ **D0** $\mu = \sigma/\sigma_{SM} = 1.2 \pm 1.1$

D0+CDF $\mu = \sigma/\sigma_{SM} = 1.95 \pm 0.75$