Search for a "boosted" Higgs with the CMS-Experiment

Bundesministerium für Bildung und Forschung

PHYSICS AT THE TERA SCALE
Helmholtz Alliance

Peter Vonhoegen
Thomas Hebbeker   Arnd Meyer

Physikalisches Institut III A, RWTH Aachen

2nd December 2010

RHEINISCH-WESTFÄLISCHE TECHNISCHE HOCHSCHULE AACHEN

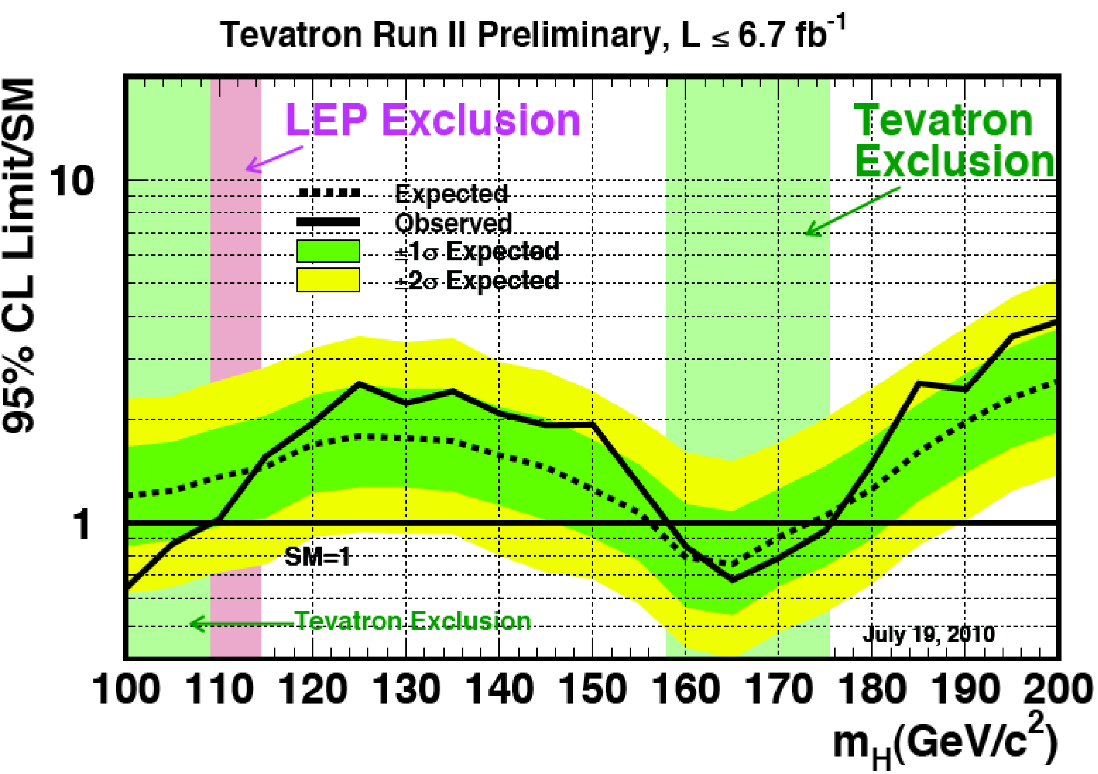
CMS

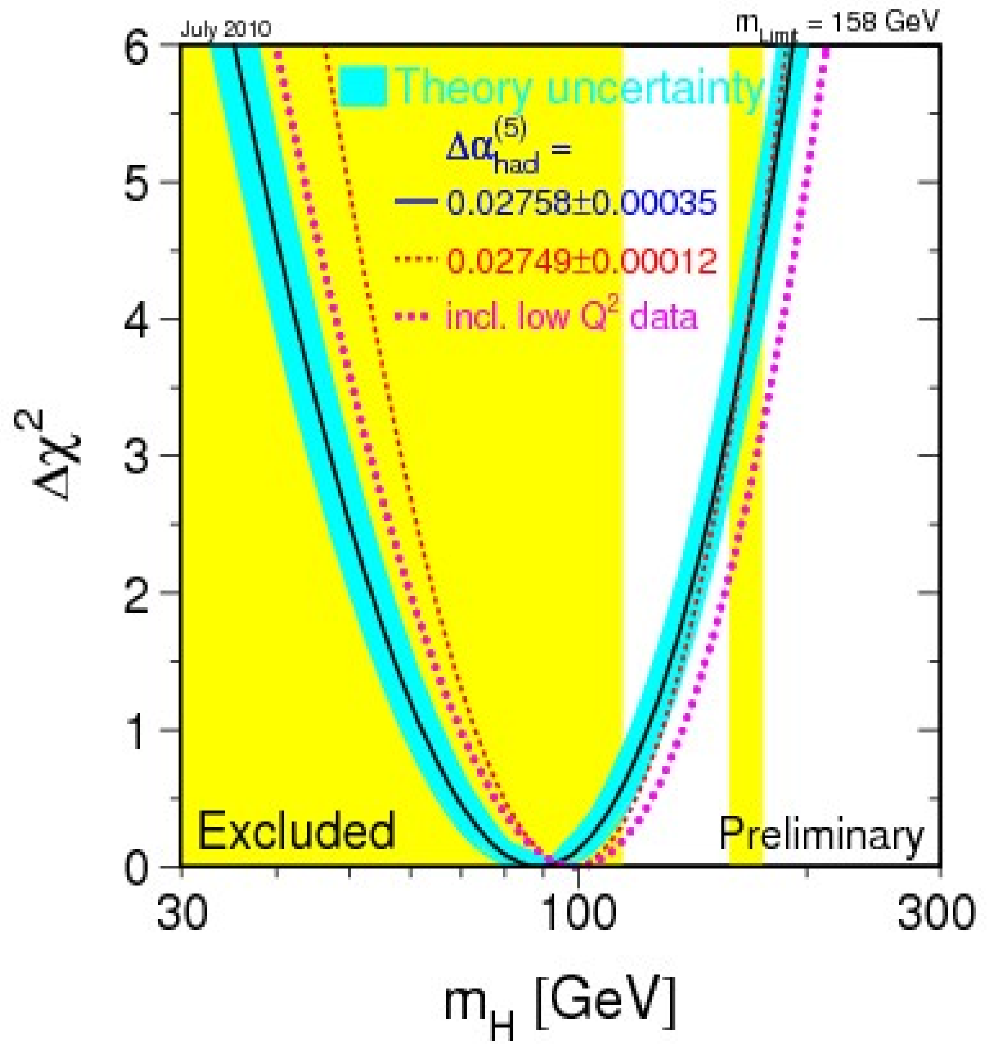
Outline

- The Higgs Boson
- The light Higgs Boson
- The “boosted Higgs”- channel
- The fat jet algorithm
- The cut based analysis
- Conclusion

The Higgs Boson

- Why do we need the Higgs?
  - To explain the masses of bosons and fermions
  - To avoid too large cross sections in W-W-scattering
- Which mass does it have?

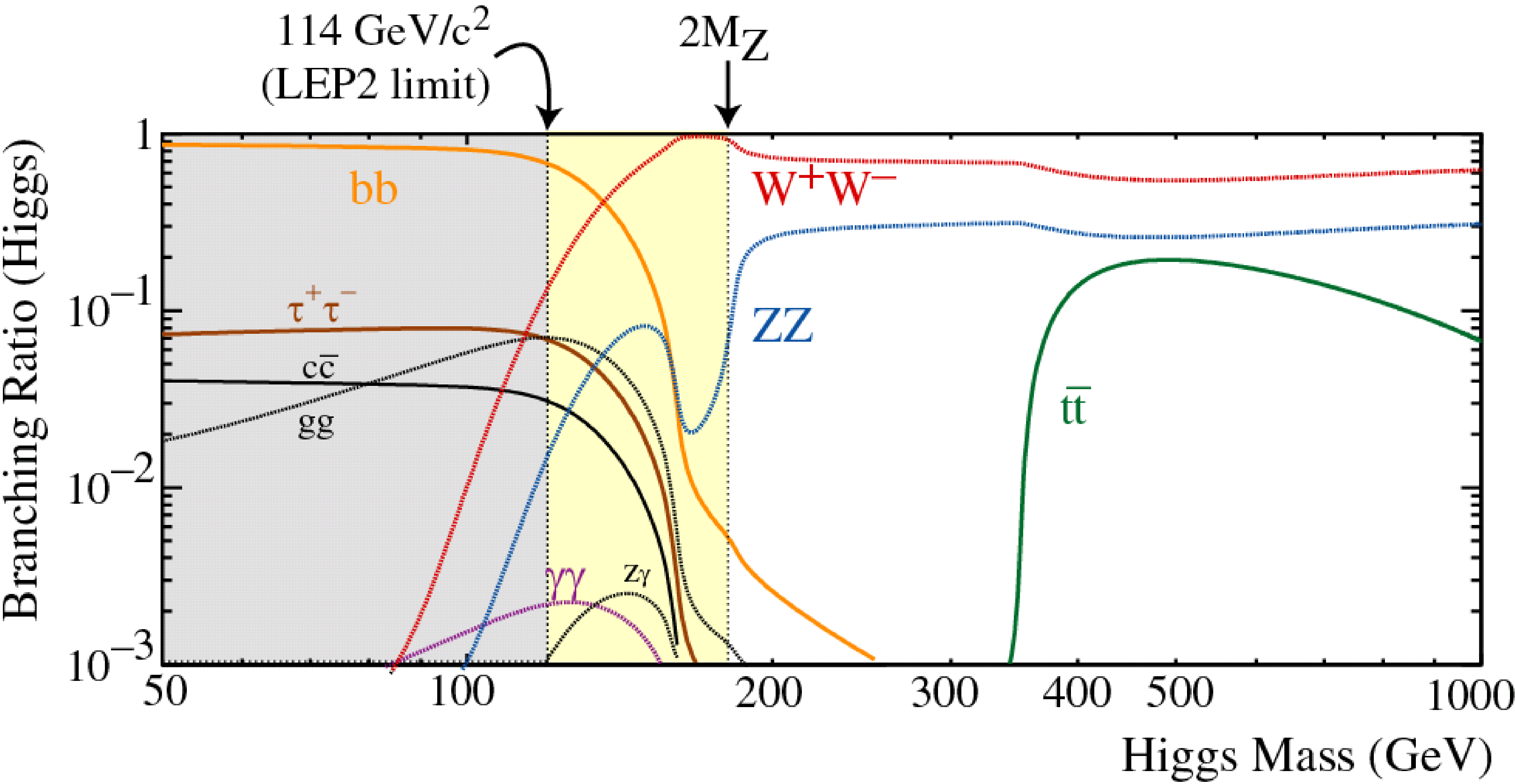




$\Rightarrow$ Most likely a light Higgs:
115 GeV < m_H < 130 GeV

The light Higgs Boson

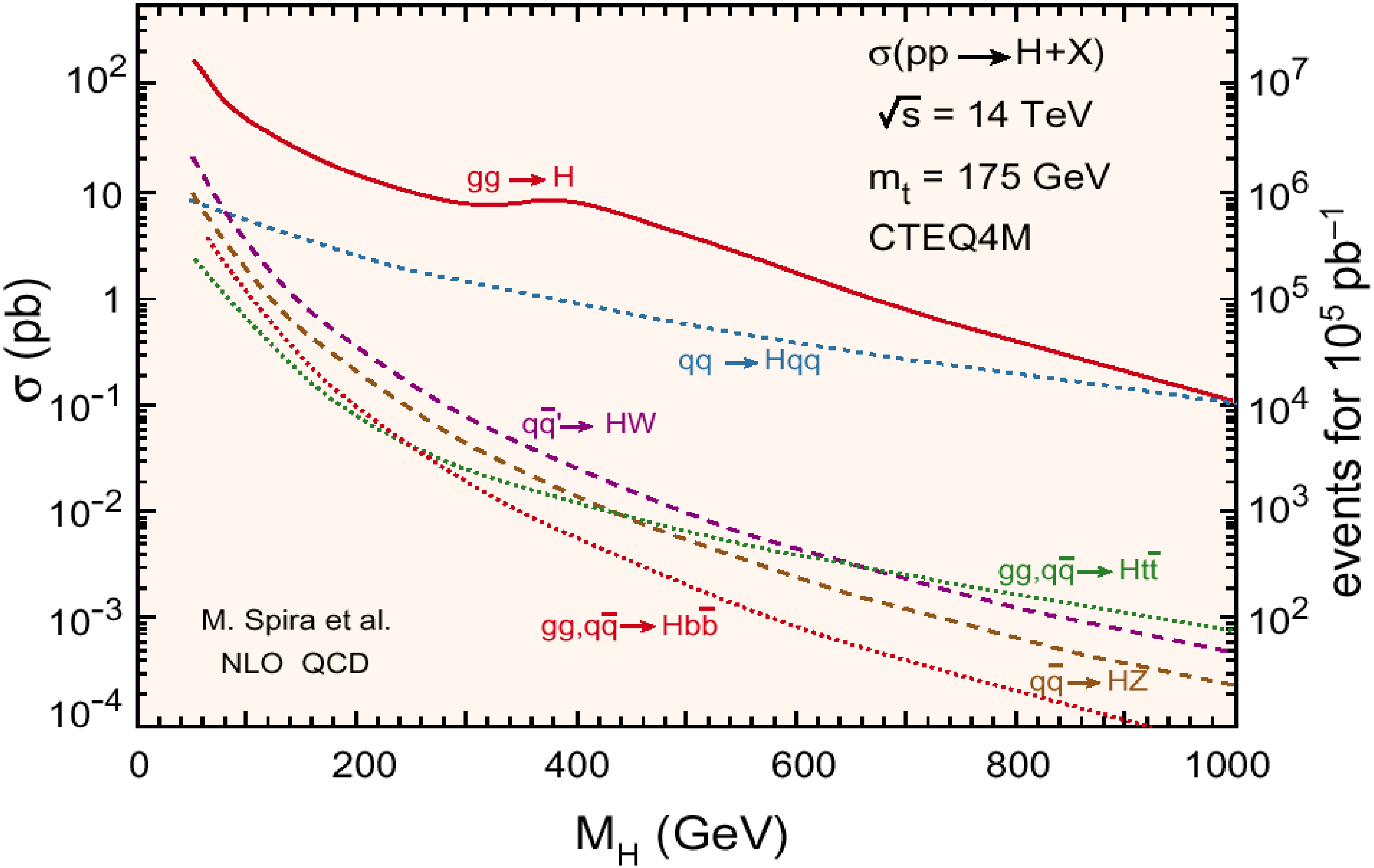
- Most likely decay channel for a light Higgs: $\mathbf{H \rightarrow b\bar{b}}$



The light Higgs Boson

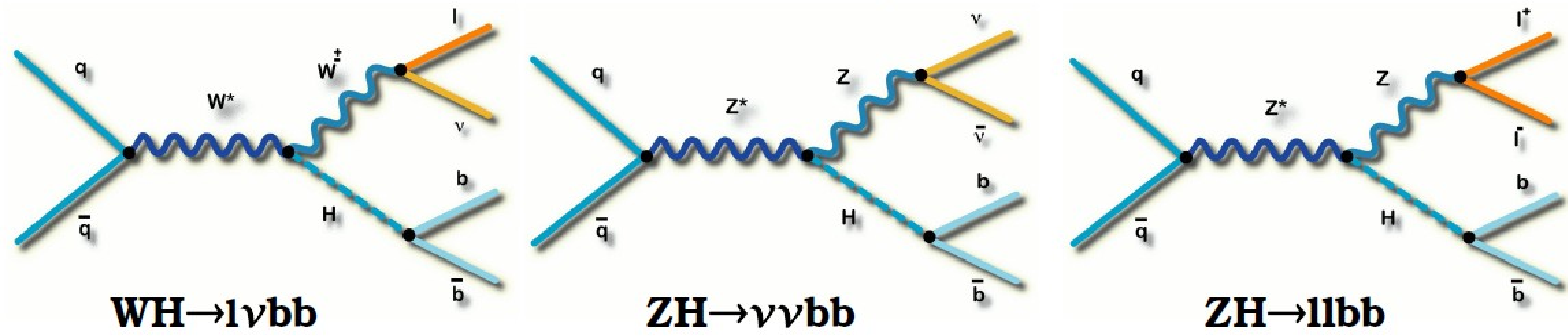
- At hadron-colliders: large QCD-backgrounds

$\Longrightarrow$ Look at associated Higgs production with a leptonically decaying vector boson to get a clear event signature: $\mathbf{q\bar{q} \to HV}$



The “boosted Higgs”-Channel

- 3 sub-channels:



- Main light Higgs search channels at TeVatron!
- But until now mostly ignored at LHC,
  because at LHC energies the QCD background seems too difficult to suppress
- New idea for a way out: look at the boosted regime!
  - H and V back to back, both with large transverse momenta

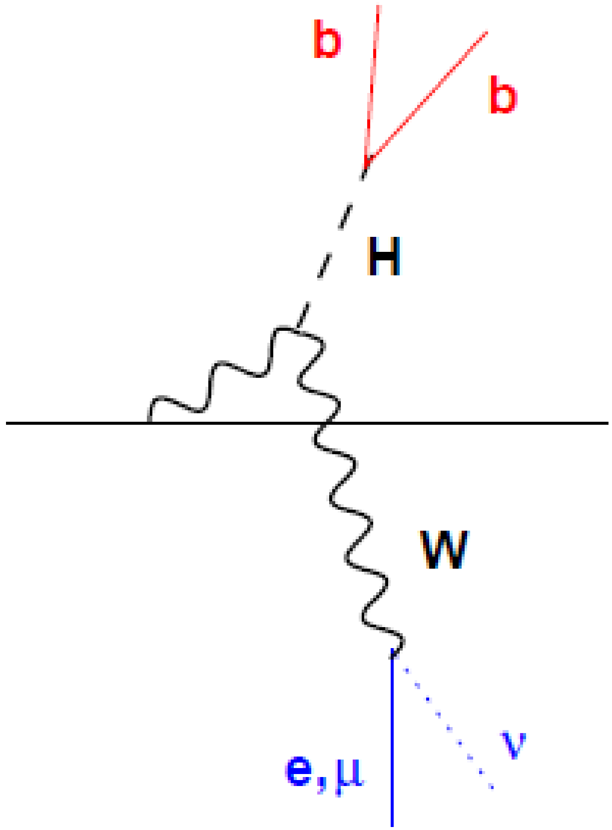
(Butterworth et al., arxiv: [0802.2470])

The "boosted Higgs"-Channel

- Advantages of a boosted Higgs:
  - Vector boson and Higgs are all central
  - Better b-tagging and jet resolution
  - Excellent background rejection (QCD, ttbar)
  - $\mathbf{Z} \rightarrow \nu\nu$ becomes visible (high MET)

$\Rightarrow$ Clean event topology

$\Rightarrow$ Good significance



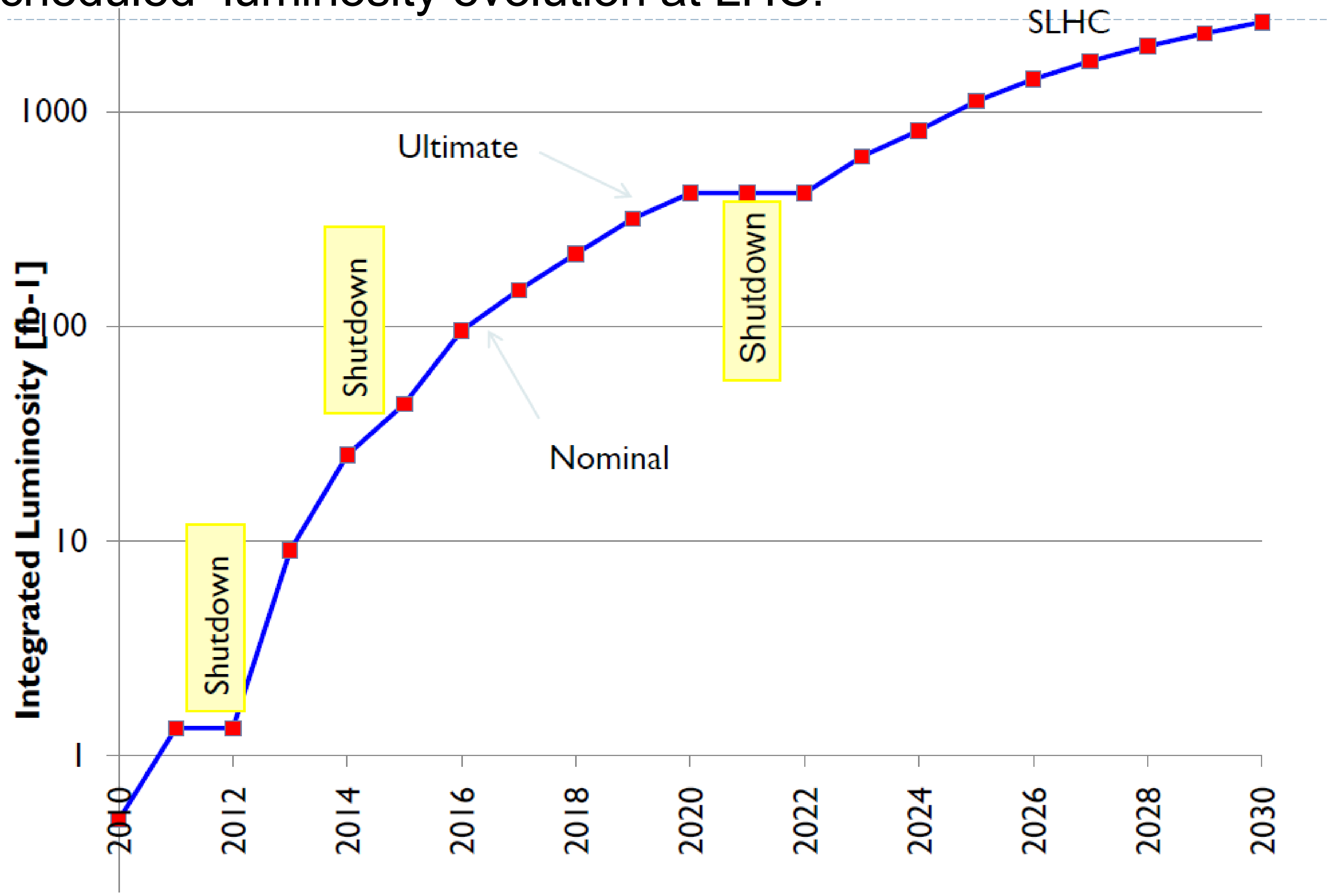
- Disadvantage:
  - Only 5 % of the produced Higgs $(\mathbf{q\bar{q} \rightarrow HV})$ have a sufficiently high transverse momentum $(\mathbf{p_t} \gtrsim \mathbf{200}\ \mathrm{GeV})$

$\Rightarrow$ need high integrated luminosity $(\approx \mathbf{30\ fb^{-1}})$

but the other light Higgs channels like H → γγ need luminosities of the same order!

The “boosted Higgs”-Channel

- Scheduled luminosity evolution at LHC:



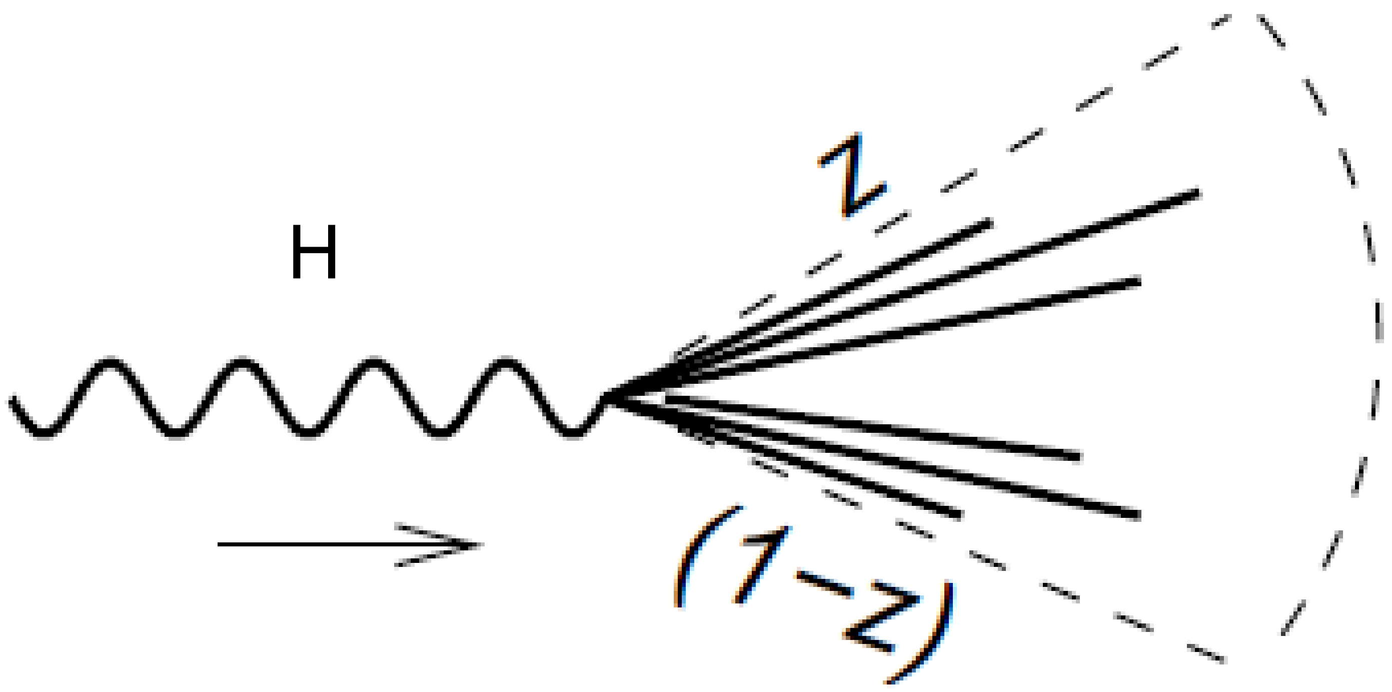
$\Rightarrow$ long term study for 14 TeV

[M. Lamont]

The “boosted Higgs”-Channel

- Apparent problem to be solved: H has high p_T

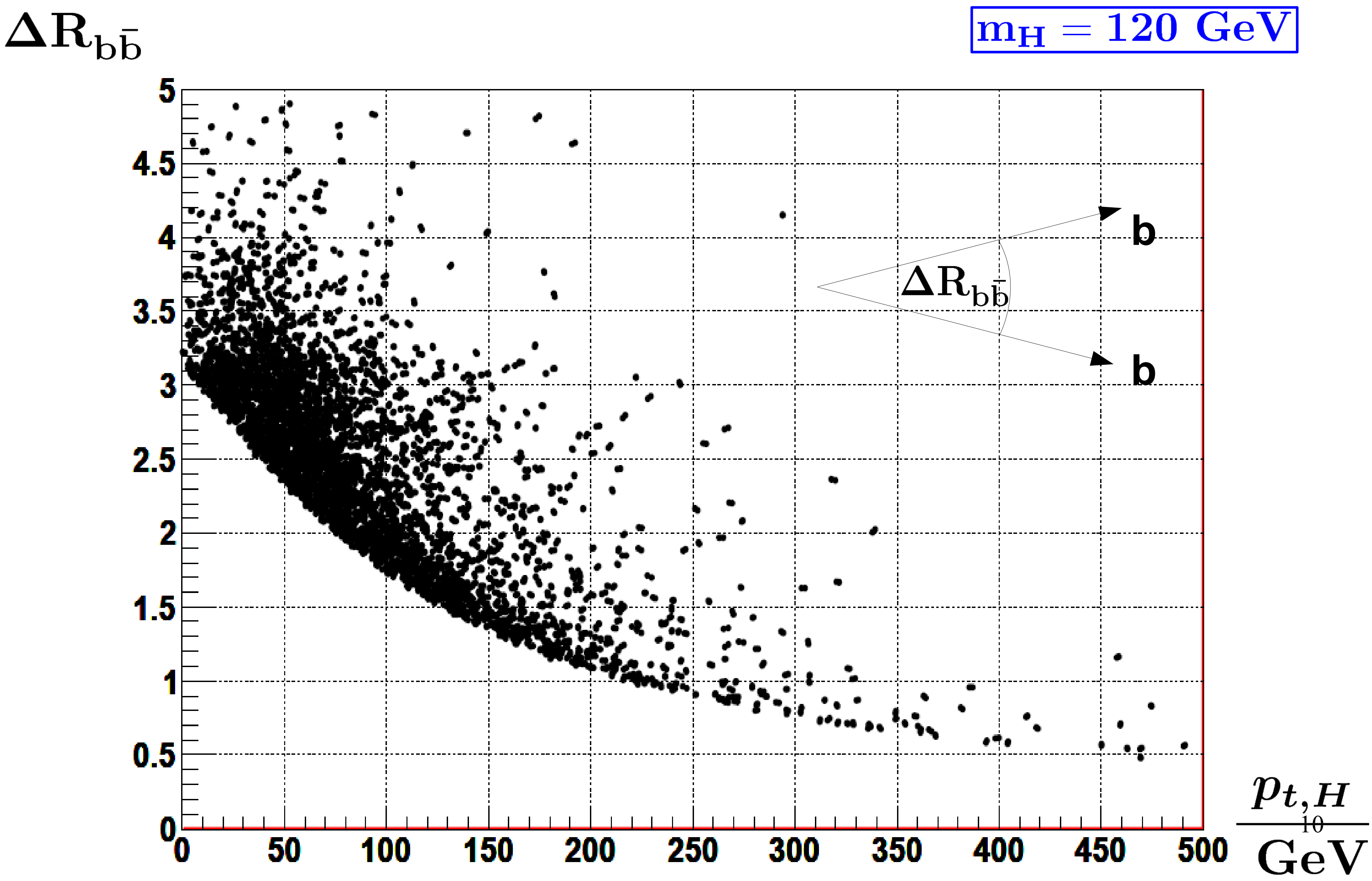
  $\Rightarrow$ b's so close together, that conventional jet reconstruction lumps them together as one single fat jet



single fat jet

$\Rightarrow$ Poor efficiency and mass resolution

The “boosted Higgs”-Channel

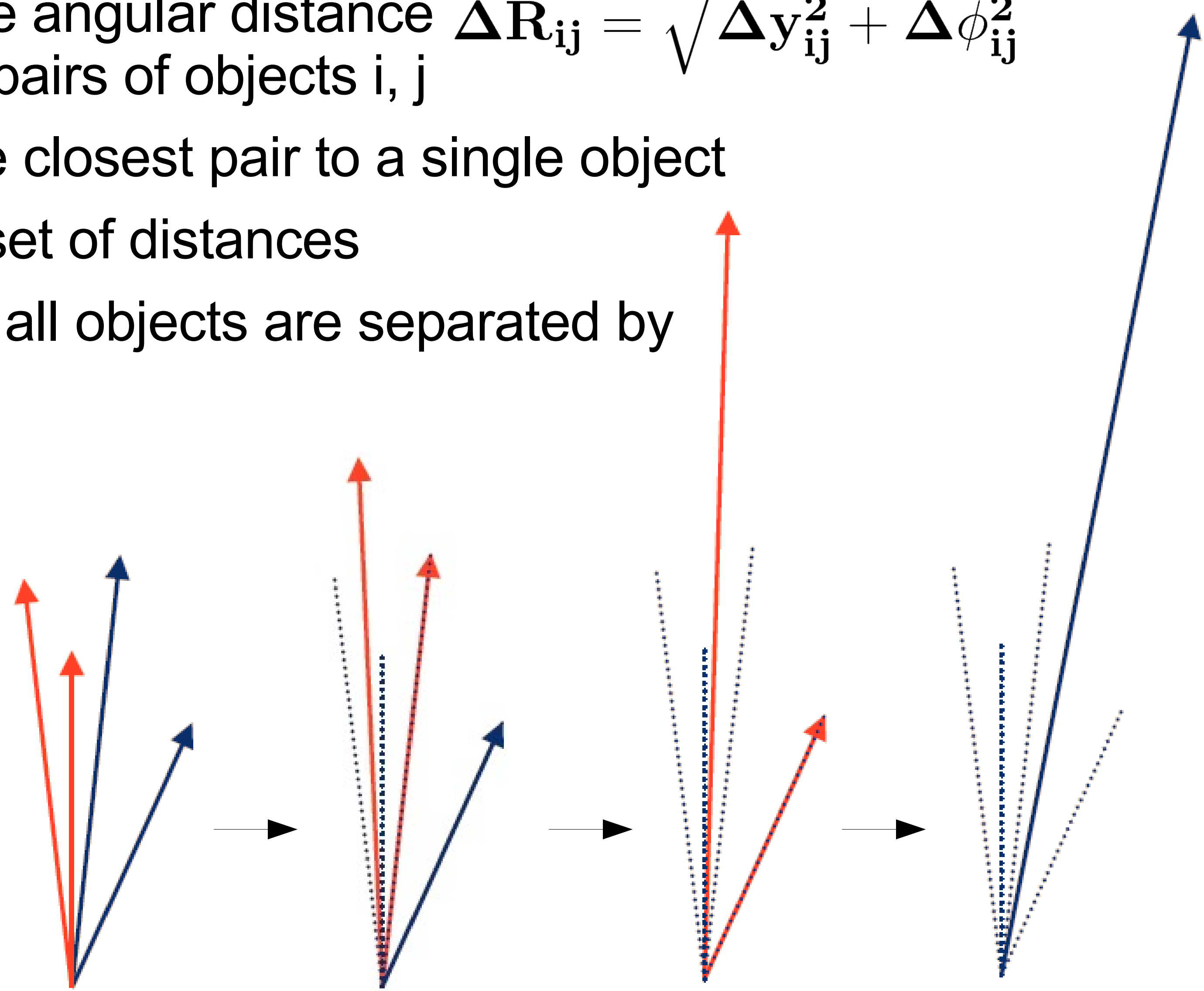


The fat jet algorithm - Clustering

- Use the iterative Cambridge/Aachen-jet-cluster-algorithm to reconstruct the fat jet:
  - Calculate the angular distance $\mathbf{\Delta R_{ij} = \sqrt{\Delta y_{ij}^2 + \Delta \phi_{ij}^2}}$ between all pairs of objects i, j
  - Combine the closest pair to a single object
  - Update the set of distances
  - Repeat until all objects are separated by $\mathbf{\Delta R_{ij} > R}$

only parameter of the algorithm

$\Longrightarrow$ hierarchical structure in angles



The fat jet algorithm - Decomposition

- given: Cambridge/Aachen-jet with radius R
- Iterative decomposition procedure to find substructure:

parameters of the algorithm

(1) Break the jet j into 2 subjets $\mathbf{j_1, j_2}\ (\mathbf{m_{j1} > m_{j2}})$
(= undo the last stage of clustering)

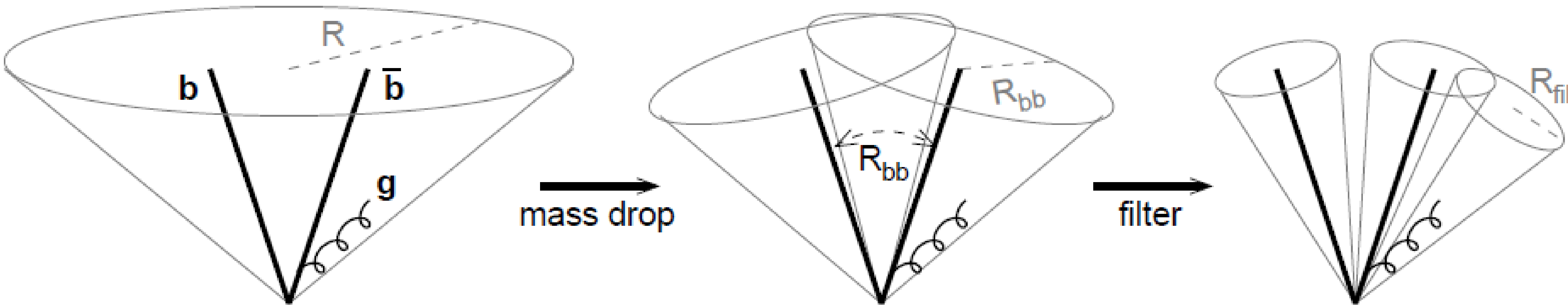
(2) If there was a significant mass drop $(\mathbf{m_{j1}} < \mu \mathbf{m_j})$
and the splitting is not too asymmetric

$$\left(\mathbf{y} = \frac{\min\left(\mathbf{p^2_{t,j1}, p^2_{t,j2}}\right)}{\mathbf{m_j^2}} \mathbf{\Delta R^2_{j1,j2}} \simeq \frac{\min\left(\mathbf{p_{t,j1}, p_{t,j2}}\right)}{\max\left(\mathbf{p_{t,j1}, p_{t,j2}}\right)} > \mathbf{y_{cut}}\right)$$

deem j to be the Higgs neighborhood and exit the loop

(3) Otherwise redefine j to be equal to j_1 and go back to (1)

The fat jet algorithm - Filtering



- Because of angular ordering, the QCD radiation will be emitted in the two cones of size R_{bb} around the b-quarks

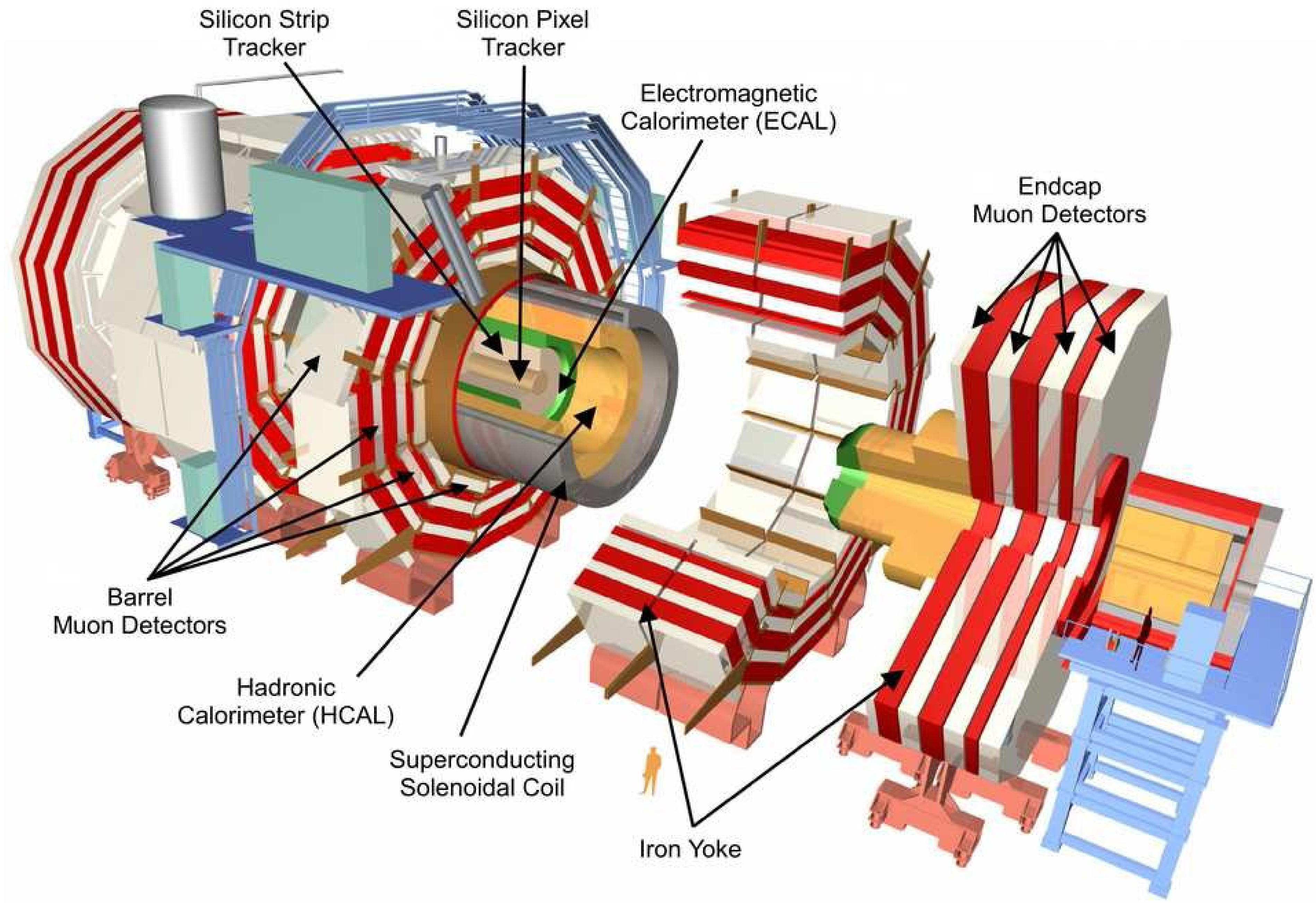
- Next step: filter the Higgs neighborhood to select the bb pair out of the bbg configuration

  - Rerun the C/A algorithm on the jet constituents, using a finer angular scale $\mathbf{R_{filt}} = \min\left(\mathbf{0.3}, \mathbf{R_{b\bar{b}}}/\mathbf{2}\right) < \mathbf{R_{b\bar{b}}}$
  - Take the 3 hardest objects (sub-jets) that appear

$\Rightarrow$ Captures b jets + $\mathcal{O}(\alpha_\mathbf{s})$ radiation & filters out underlying events $(\sim \mathbf{R^4_{b\bar{b}}})$ & pile-up

$\Rightarrow$ Improves mass and angular resolution (and with it b-tagging!)

The Analysis with CMS



Setup for the analysis

- Implementation of the “Boosted Higgs Algorithm” with help of Carsten Magass
  [Use of the FastJet Package by M. Cacciari, G. Salam and G. Soyez, http://www.lpthe.jussieu.fr/~salam/fastjet/]
- Mass production of PYTHIA Monte Carlos:
  - Signal (3x 450.000 events, FULLSIM, m_H = 115/120/130 GeV)
  - ZZ, WW, WZ with generator cut: $\mathbf{p_{T,V} > 100GeV}$
    (3 x 500.000 events, FASTSIM)
  - ttbar (10.000.000 events, FASTSIM)
- Mass production of SHERPA Monte Carlos with help of Metin Ata and Markus Merschmeyer:
  - Z + Jets with generator cut: $\mathbf{p_{T,Z} > 100GeV}$
    (2 x 1.000.000 events, FASTSIM)
  - W + Jets with generator cut: $\mathbf{p_{T,W} > 100GeV}$
    (5.000.000 events, FASTSIM)

The cut based analysis

- Cutflow:

high-p_T- MET-/electron-/muon- Trigger

Higgs Selection

[0] at least **1 fat jet**

[1] ...with **p_T > 200 GeV**

[2] ...and **|eta| < 2.5**

[3] ...where the **hardest subjet** is **b-tagged**

[4] ...and where the **second-hardest** subjet is **b-tagged**

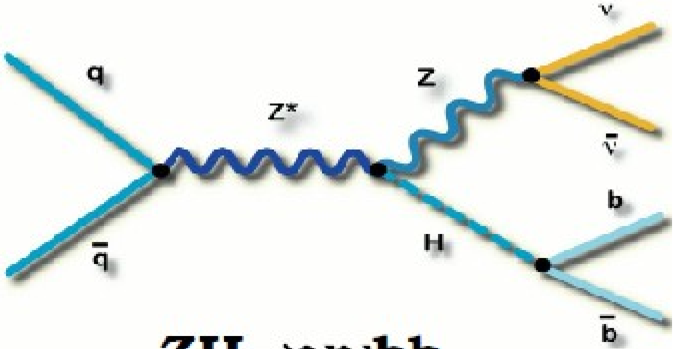
[5] **exactly** 1 such fat jet

[6] **no additional Jets**

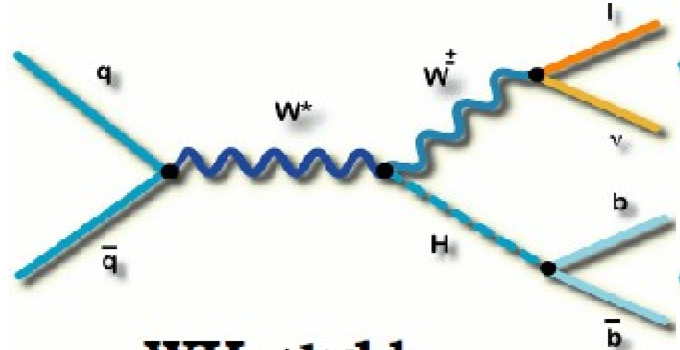
[7] **at most 2 same type leptons**

[8] **# leptons = 0** ← orthogonal → [8] **# leptons = 1** ← orthogonal → [8] **# leptons = 2**

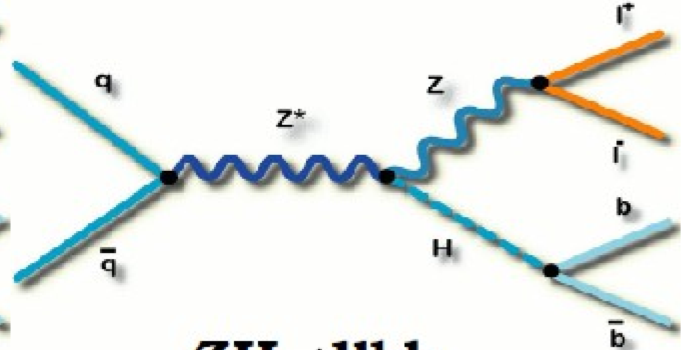
The cut based analysis



ZH→ννbb



WH→lνbb



ZH→llbb

ZH→ννbb:

- [8] **# leptons = 0**
- [9] **MET > 180 GeV**
- [10] **Sign(MET) > 11**

$$Z_0 = \frac{s_0}{\sqrt{b_0 + \sigma^2_{b_0}}}$$

WH→lνbb:

- [8] **# leptons = 1**
- [9] **$p_{T,lep}$ > 100 GeV**
- [10] **MET > 45 GeV**
- [11] **$p_{T,W}$ > 200 GeV**
- [12] **|ΔΦ (W,H)| > 3.0**

$$Z_1 = \frac{s_1}{\sqrt{b_1 + \sigma^2_{b_1}}}$$

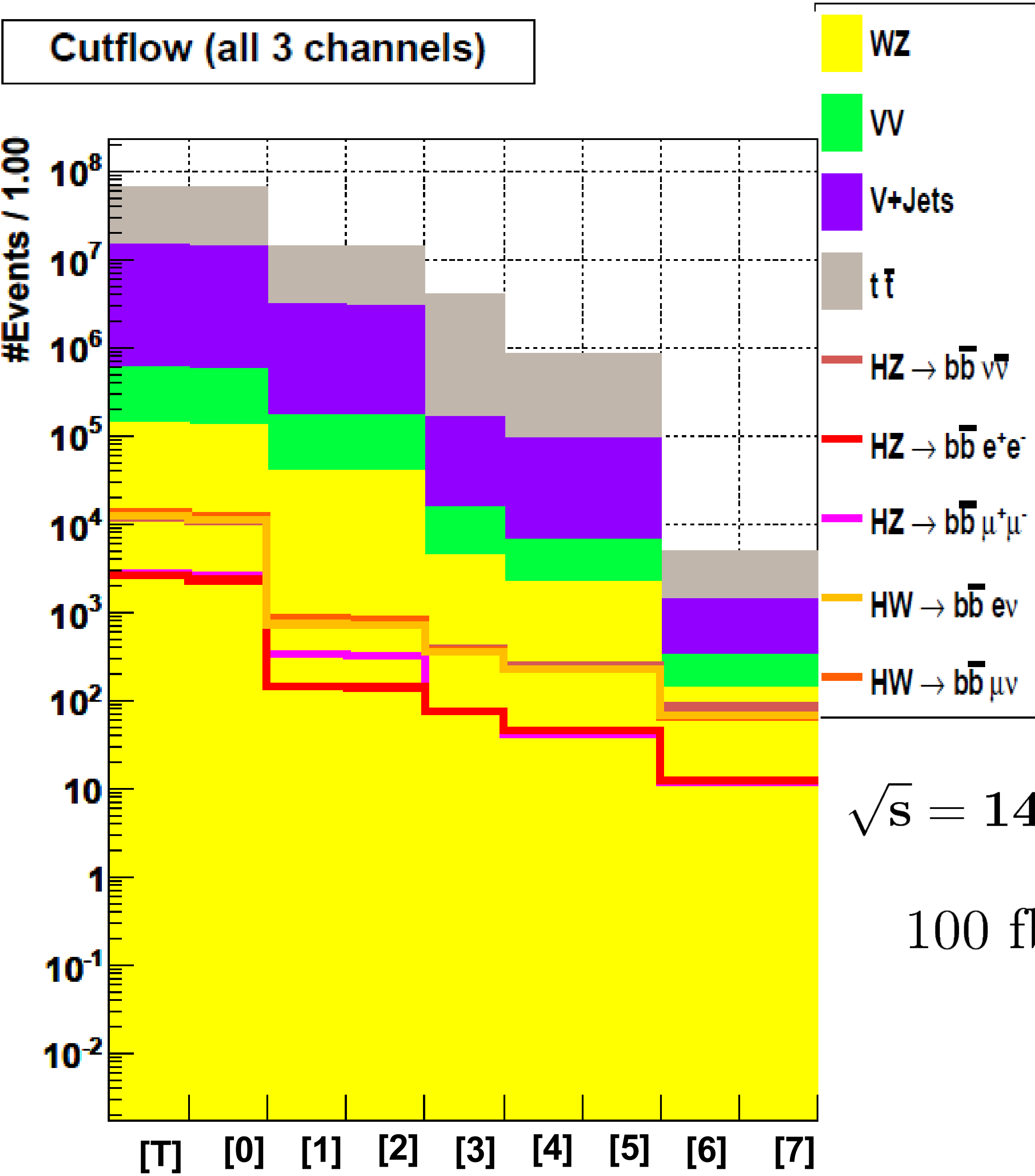
ZH→llbb:

- [8] **# leptons = 2**
- [9] **$p_{T,lep1}$ > 110 GeV**
- [10] **$p_{T,lep2}$ > 20 GeV**
- [11] **75 GeV<m_Z<105 GeV**
- [12] **$p_{T,Z}$ > 200 GeV**

$$Z_2 = \frac{s_2}{\sqrt{b_2 + \sigma^2_{b_2}}}$$

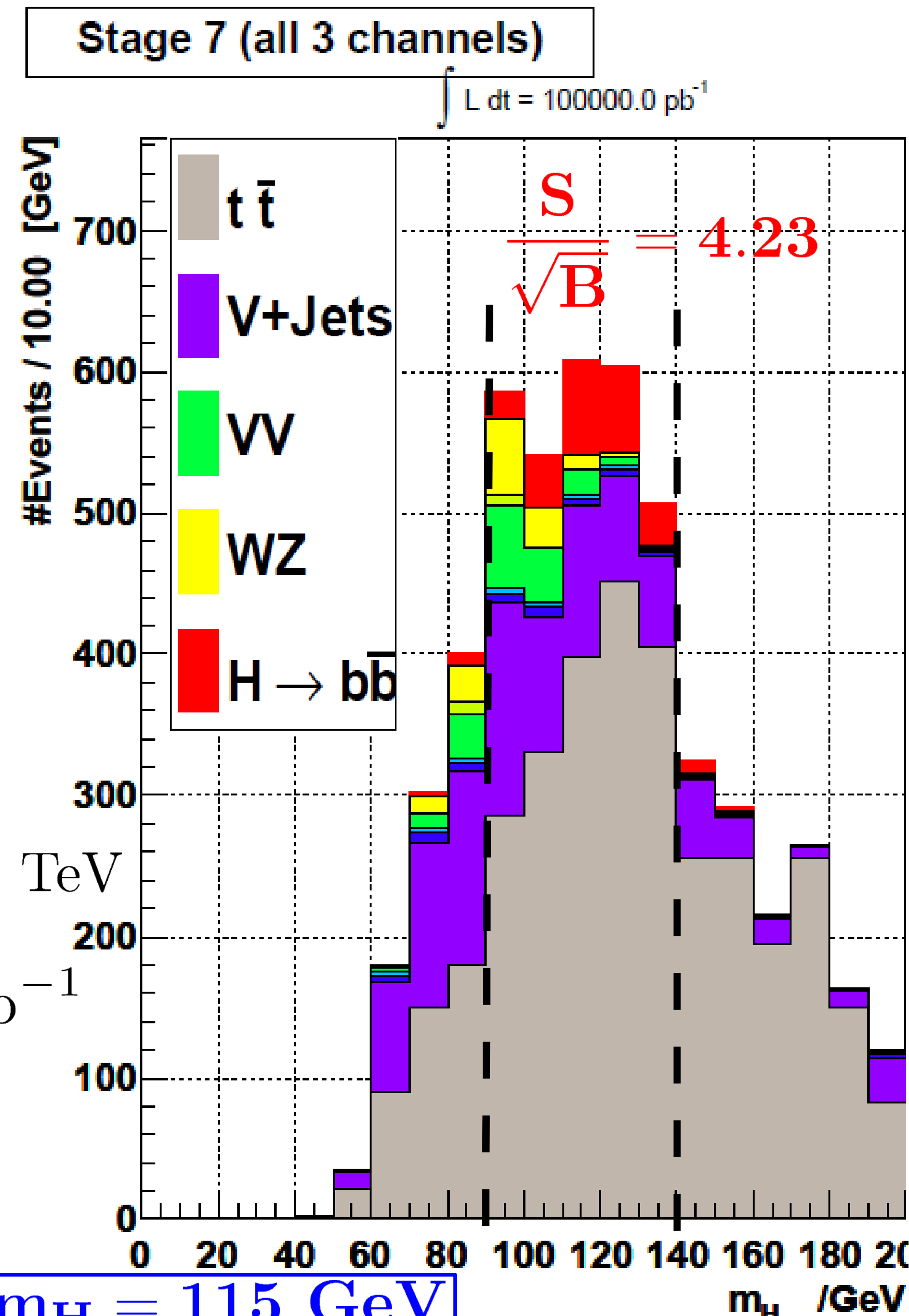
$$\Rightarrow \text{total Significance } Z = \sqrt{Z_0^2 + Z_1^2 + Z_2^2}$$

The cut based analysis (Higgs Selection)



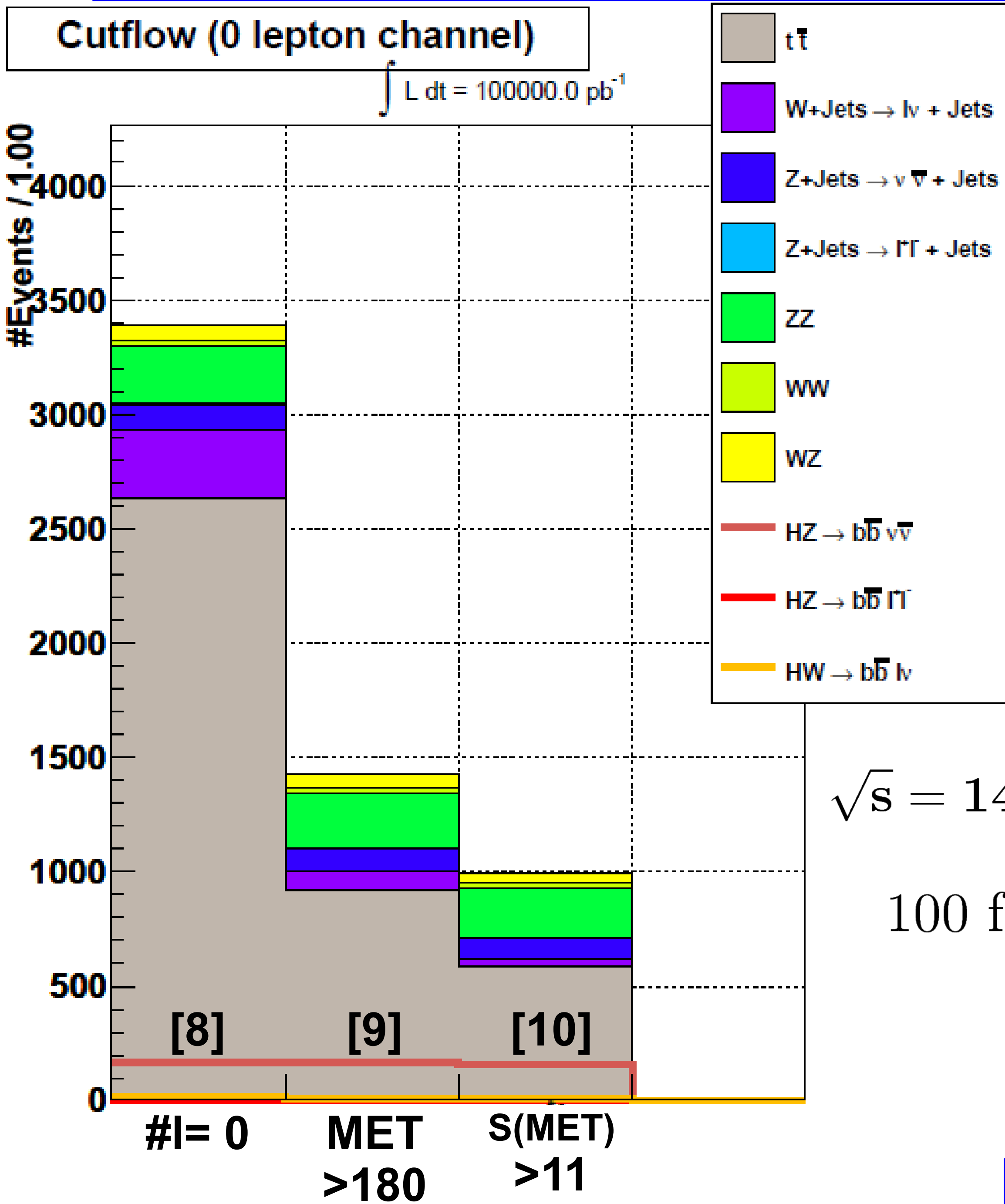
$\sqrt{s} = 14$ TeV

$100\ \mathrm{fb}^{-1}$



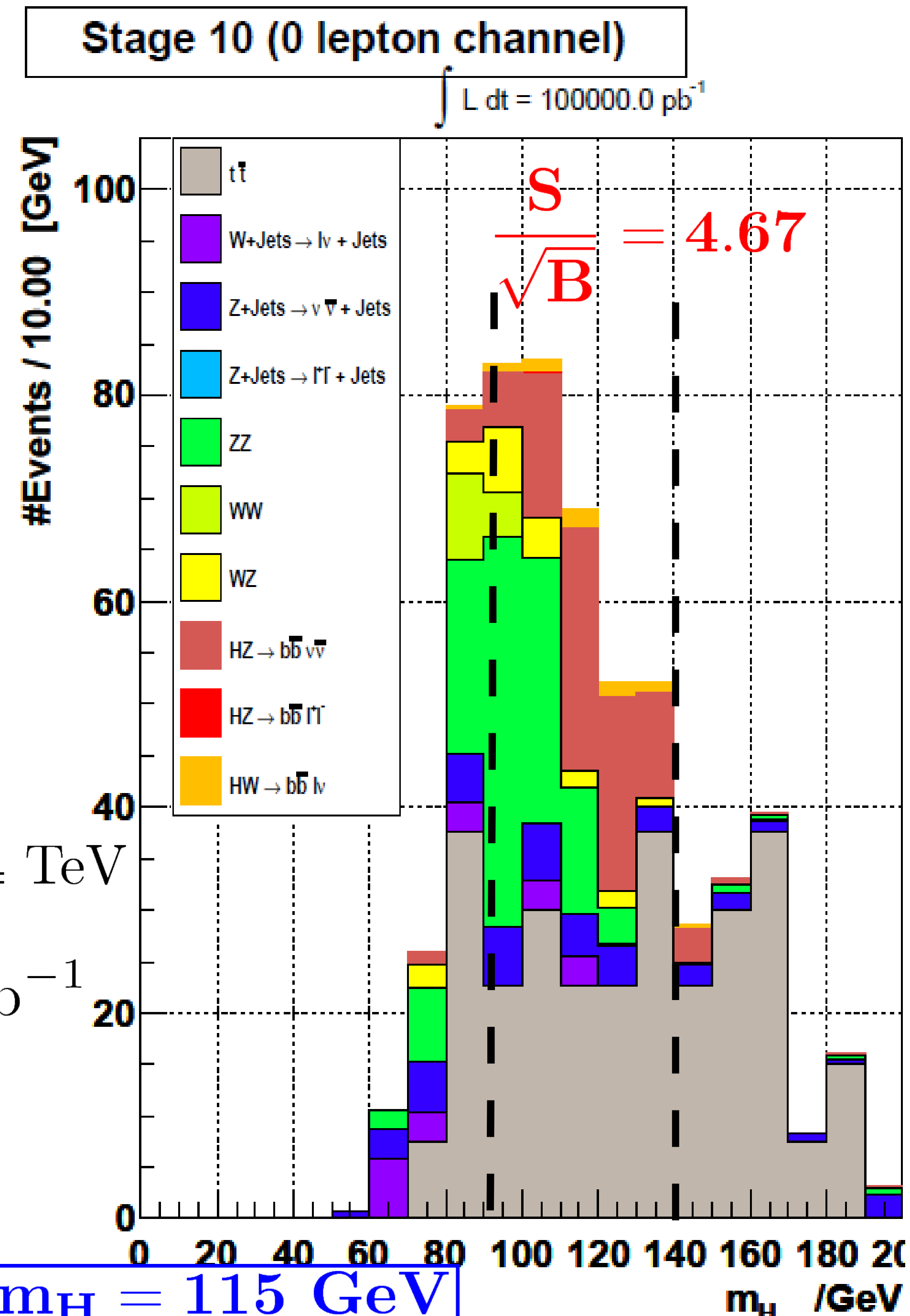
$m_H = 115$ GeV

The cut based analysis $(\mathrm{HZ} \to \mathrm{b\bar{b}}\nu\bar{\nu})$



$\sqrt{s} = 14$ TeV

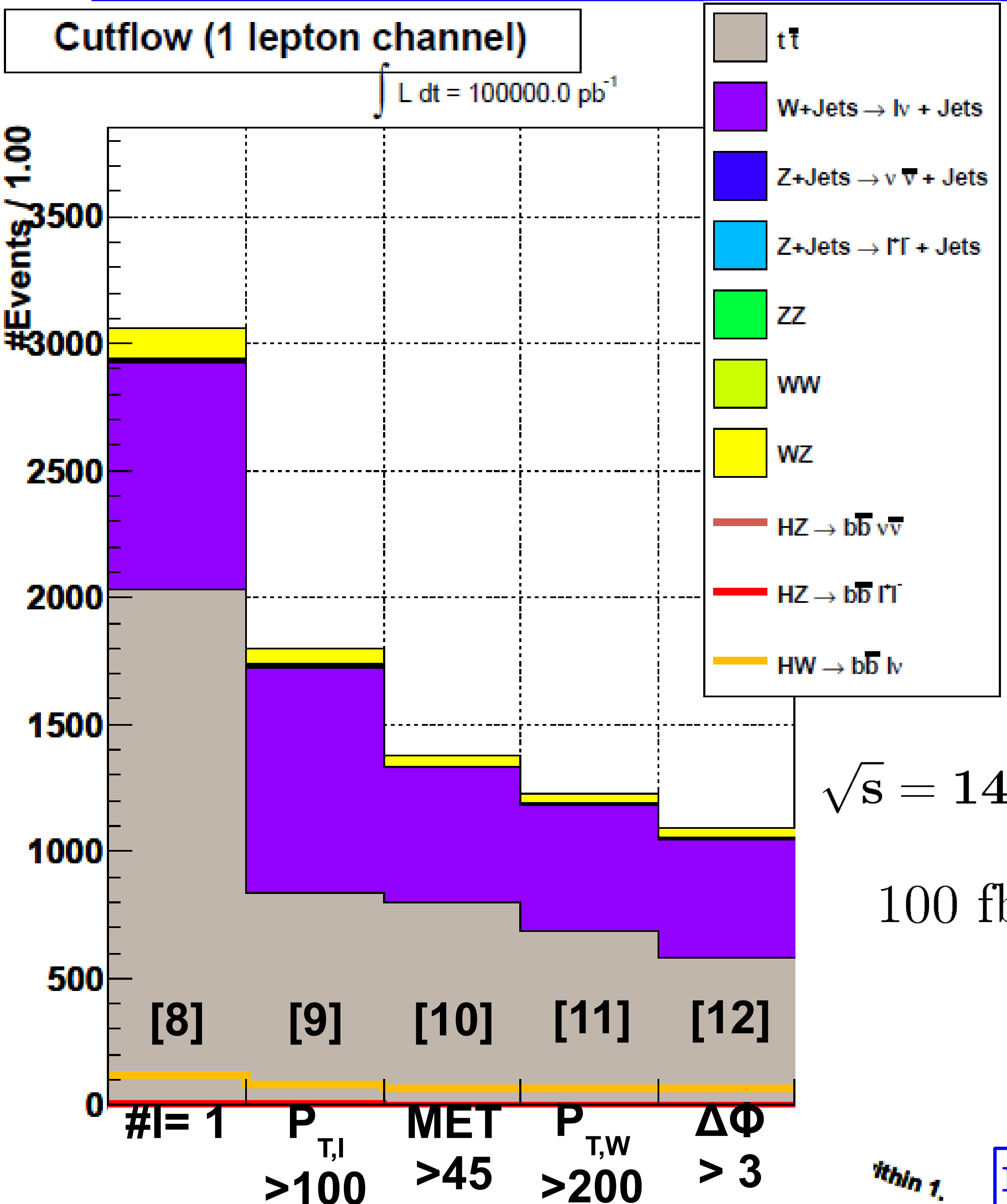
$100\ \mathrm{fb}^{-1}$



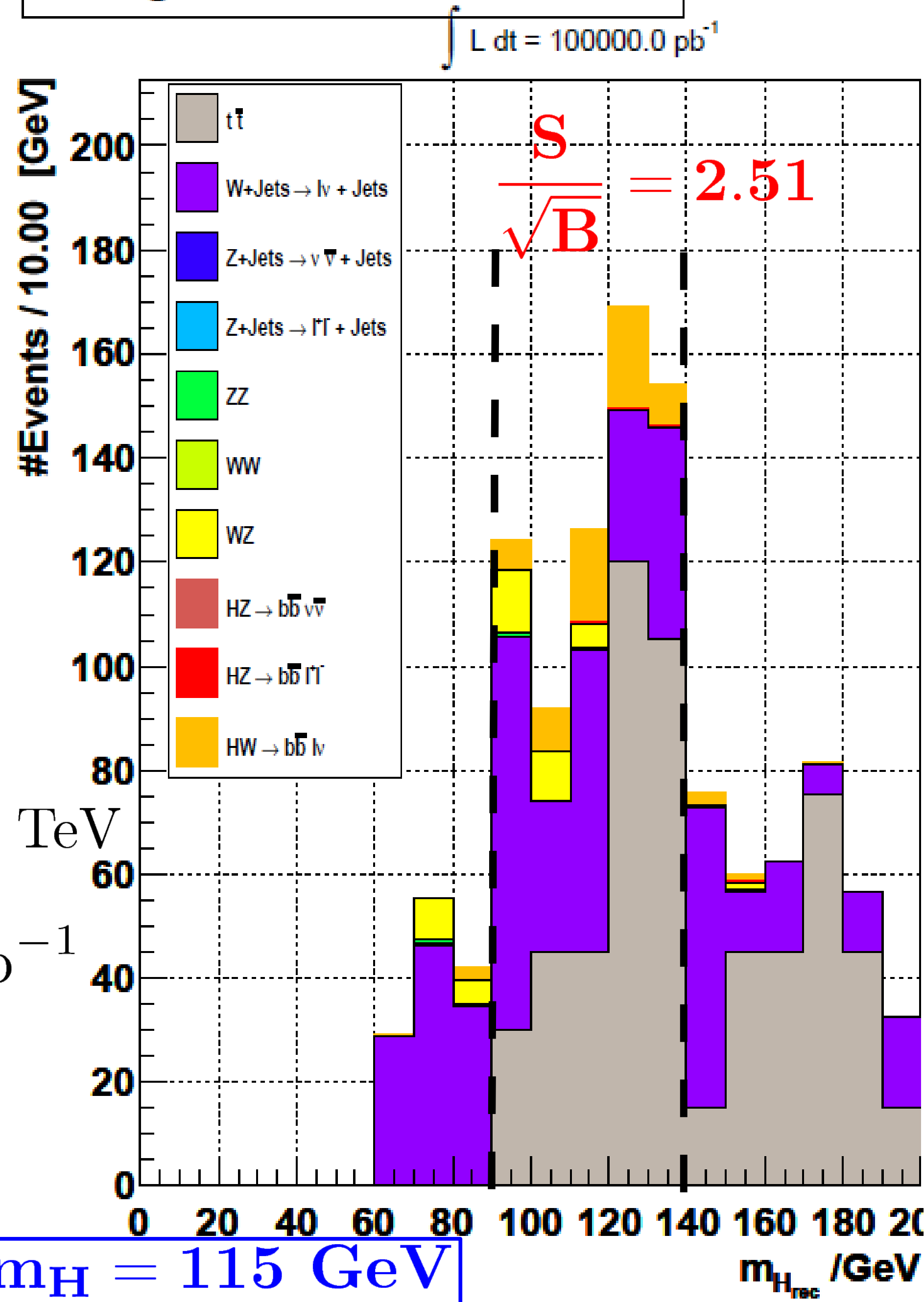
$\boxed{m_H = 115\ \mathrm{GeV}}$

The cut based analysis $(\mathbf{HW} \rightarrow \mathbf{b\bar{b}l}\nu)$

Cutflow (1 lepton channel)



Stage 12 (1 lepton channel)

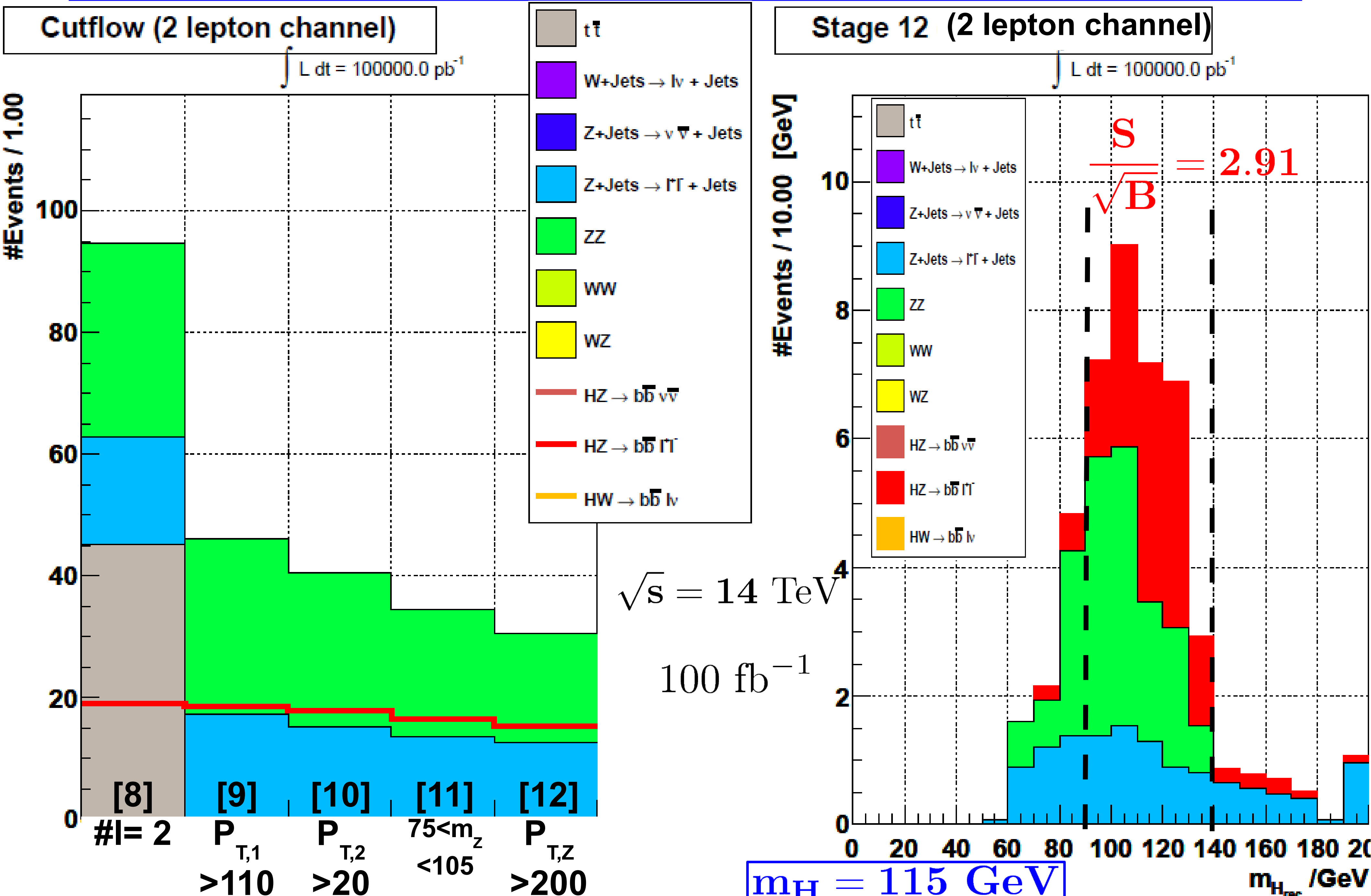


$\sqrt{s} = 14$ TeV

$100\ \mathrm{fb}^{-1}$

$m_H = 115$ GeV

The cut based analysis $(\mathrm{HZ} \rightarrow \mathrm{b\bar{b}l^{+}l^{-}})$



$$\frac{\mathrm{S}}{\sqrt{\mathrm{B}}} = 2.91$$

$$\sqrt{\mathrm{s}} = 14\ \mathrm{TeV}$$

$$100\ \mathrm{fb}^{-1}$$

$$\mathrm{m_H} = 115\ \mathrm{GeV}$$

Results

- Systematic uncertainties on the background to consider:
  - Uncertainty on the Jet Energy Scale (3%)

    $\rightarrow$ move the scale 3% up and 3% down and look how this changes b (have also to regard the consequences on MET)
  - Uncertainty on the used background cross sections (5%)
  - Uncertainty on the b-tagging efficiency (2% per jet)

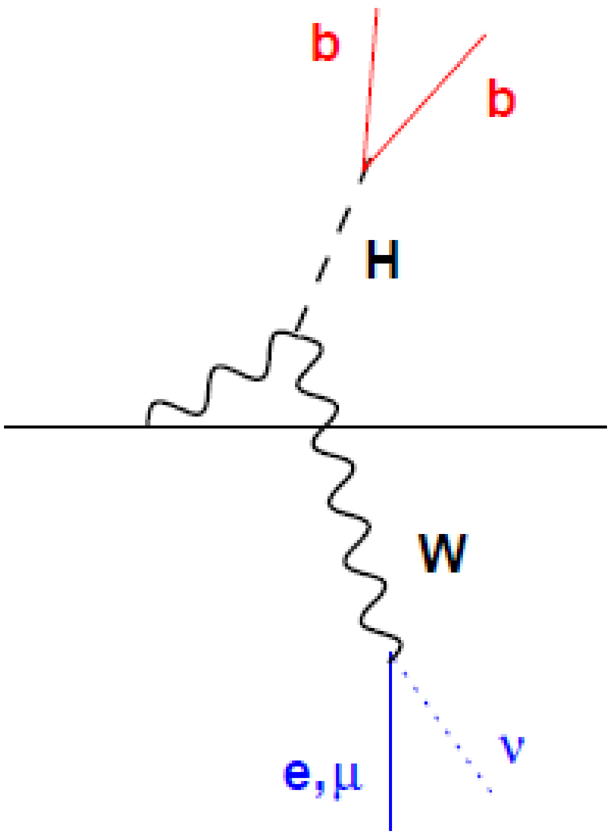
	m_H = 115 GeV		m_H = 120 GeV		m_H = 130 GeV	
	S/sqrt(B)	Z	S/sqrt(B)	Z	S/sqrt(B)	Z
0 lepton	4.67	3.18	4.03	2.79	2.59	1.65
1 leptons	2.51	1.40	2.18	1.24	1.49	0.81
2 leptons	2.91	2.67	2.28	2.10	1.62	1.59
combined	6.05	4.38	5.12	3.71	3.40	2.43

$$\mathbf{Z_i} = \frac{\mathbf{s_i}}{\sqrt{\mathbf{b_i} + \sigma^2_{\mathbf{b_i}}}}$$

$$Z = \sqrt{Z_0^2 + Z_1^2 + Z_2^2}$$

Summary

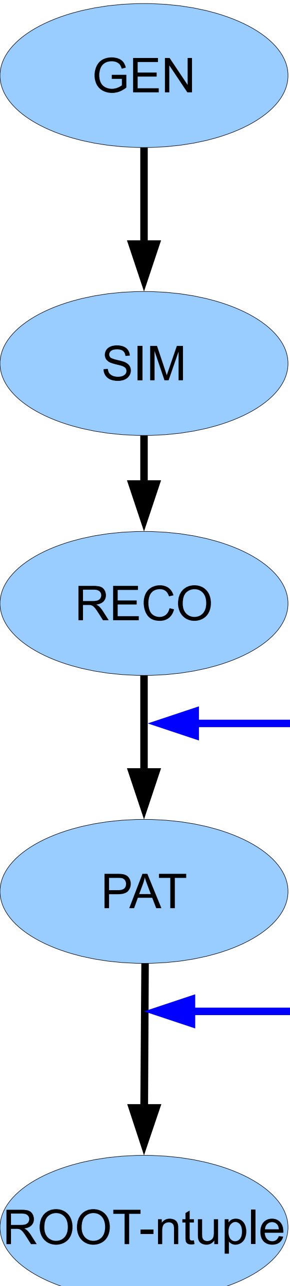
- Looking for a light Higgs boson (115-120 GeV):
  - produced together with vector boson
  - boosted regime
  - $H \to bb$, $V \to$ leptons
- Hopeless case recovered as promising way
- Main instrument: the fat jet algorithm
- Combined significance for a scenario of 100 fb^{-1} and m_H= 115 GeV: Z = 4.38



Thank You!

BackUp Slides

The Implementation in CMS



Generators (PYTHIA, Madgraph, SHERPA, ...) produce a lot of events based on the theoretical probabilities of the standard model. Input: Center of mass energy, parton distribution function (pdf), channel

The Detector-Impact (Hits in Tracker, Calorimeters etc.) of the generated particles is **sim**ulated.

Particle tracks and related energy clusters are **reco**nstructed. (The first step, that is used on both: Monte-Carlos and Data)

Boosted-Higgs-algorithm (all Parameters set in Config-File)

“Physics Analysis Toolkit”. Physical objects like photons, electrons, muons, jets, met etc. are reconstructed

AC-Susy-Filler (implemented in CMSSW by Carsten Magass)

The properties of simple analysis objects (e.g. transverse momentum of the n^{th} electron in the m^{th} event, or number of fatjets in the x^{th} event) are available in root trees.