Study of double Higgs production at 3 TeV (and 1.4 TeV)

Rosa Simoniello

CLICdp collaboration meeting, 2-3 June 2015, CERN

Many thanks to Philipp Roloff

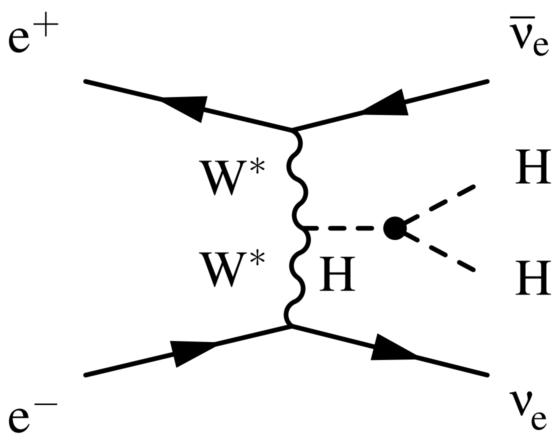
Outline

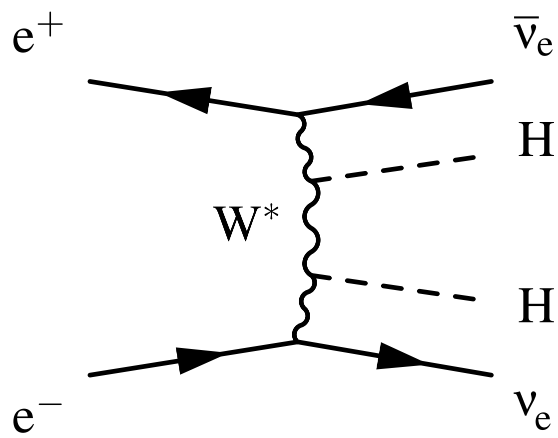
- Brief introduction
- Analysis strategy
  - Focus on object identification, in particular jets

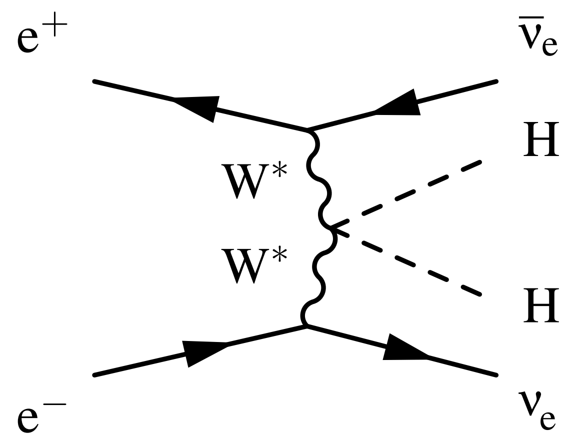
NOTE: *CLIC_SiD detector used*

→ Fix to correct PFOs selection is run

Introduction







Trilinear Higgs self coupling, g_{HHH}

Quartic coupling, g_{HHWW}

- Measurement of the Higgs self-coupling g_{HHH} at 3 TeV (and 1.4 TeV)
- Higgs production in WW fusion → it gives also diagrams with quartic coupling g_{HHWW}
- At the moment, focus on HH->bbbb channel
- Plan to include HH->WWWW channel (as uncorrelated analysis)

Analysis strategy

- Workflow:
  - Rejection of beam-induced background: TightSelectedPFOs ✓
  - Lepton (e, μ) veto ✓
  - Tau veto ✓
  - Jet opt for mass reco ✓
  - b-tag ✓
  - Selection and MVA analysis

Object identification optimisation

For the moment all studies have been done only at 3 TeV, but optimisation also at 1.4TeV is planned

Lepton and Tau identification

- Isolated lepton and tau identification is applied as a veto in the analysis to reject WWnunu backgrounds (see backups)
- ID working point optimised on WWnunu sample at 3 TeV

Requirements for isolated leptons:

- d0<0.03, z0<0.04, R0<0.06
- 0.05<Rcal<0.25 || Rcal>0.9
- cosθ=0.995 (scan in: 0.990, 0.995, 0.999)
- Polynomial isolation: A=0.0, B=5.7, C=-50.
- Etrack > 15 GeV

- **Fake reco in WWnunu sample ~1%**
- **Rejected 30% of total WWnunu evts (39% of WW evts have leptons)**
- **Rejected 18% of Higgs evts not in bb**

Requirements for taus:

- pT = 1 GeV: scan in (1, 2)
- pTseed = 10 GeV: scan in (5, 10)
- Search cone = 0.03: scan in (0.02, 0.03, 0.05, 0.07 ,0.09, 0.11)
- Isolation cone = 0.3: scan in (0.2,0.3,0.4)
- Isolation energy = 3 GeV: scan in (3.,5.,10.)
- Invariant mass = 2 GeV: scan in (1.5,2.,2.5)

- **Purity of ~51% with eff of ~25%**
- **Rejected 0.1% of signal evts**

Jet algorithms

- Choice of the best jet algorithm and relative parameters:
  - **Exclusive longitudinally invariant k_t:**
    Clustering in 3, 4, 5, 6 jets, scan in R
    → in this slides shown: 4 and 5 jets
  - **Exclusive Valencia algorithm (VLC) with $\beta = 1$:**
    Clustering in 4, 5 jets, scan in R, γ = 1.0, 0.8, 0.5, 1.3
    → in this slides shown: 4 and 5 jets with γ = 1.0
  - **Inclusive anti-k_t:**
    Minimum p_T = 4, 7, 10, 15 GeV, scan in R
    → in this slides shown best case: p_T = 4 GeV
  - **Inclusive SISCone:**
    Minimum p_T = 4, 7, 10, 15 GeV, scan in R
    → in this slides shown best case: p_T = 4 GeV

Reminder

Longitudinal invariant algorithms:

$$d_{ij} = min(p_{T,i}^{2n}, p_{T,i}^{2n})\Delta R_{ij}^{2n}/R^2$$

$$d_{iB} = p_{T,i}^{2n}$$

$$\text{for } n = 1 \ \rightarrow k_t, \text{for } n = -1 \ \rightarrow anti\text{-}k_t$$

Valencia algorithm:

$$d_{ij} = min(E_i^{2\beta}, E_i^{2\beta})(1 - cos\theta_{ij})/R^2$$

$$d_{iB} = E^{2\beta} sin^{2\gamma}\theta_{iB}$$

$$\text{for } \beta = \gamma = 1 \ \rightarrow d_{iB} = p_{T,i}^2$$

Reference paper: https://inspirehep.net/record/1291037?ln=en

Talk at the CLIC workshop: https://indico.cern.ch/event/336335/session/1/contribution/176/material/slides/0.pdf

Info

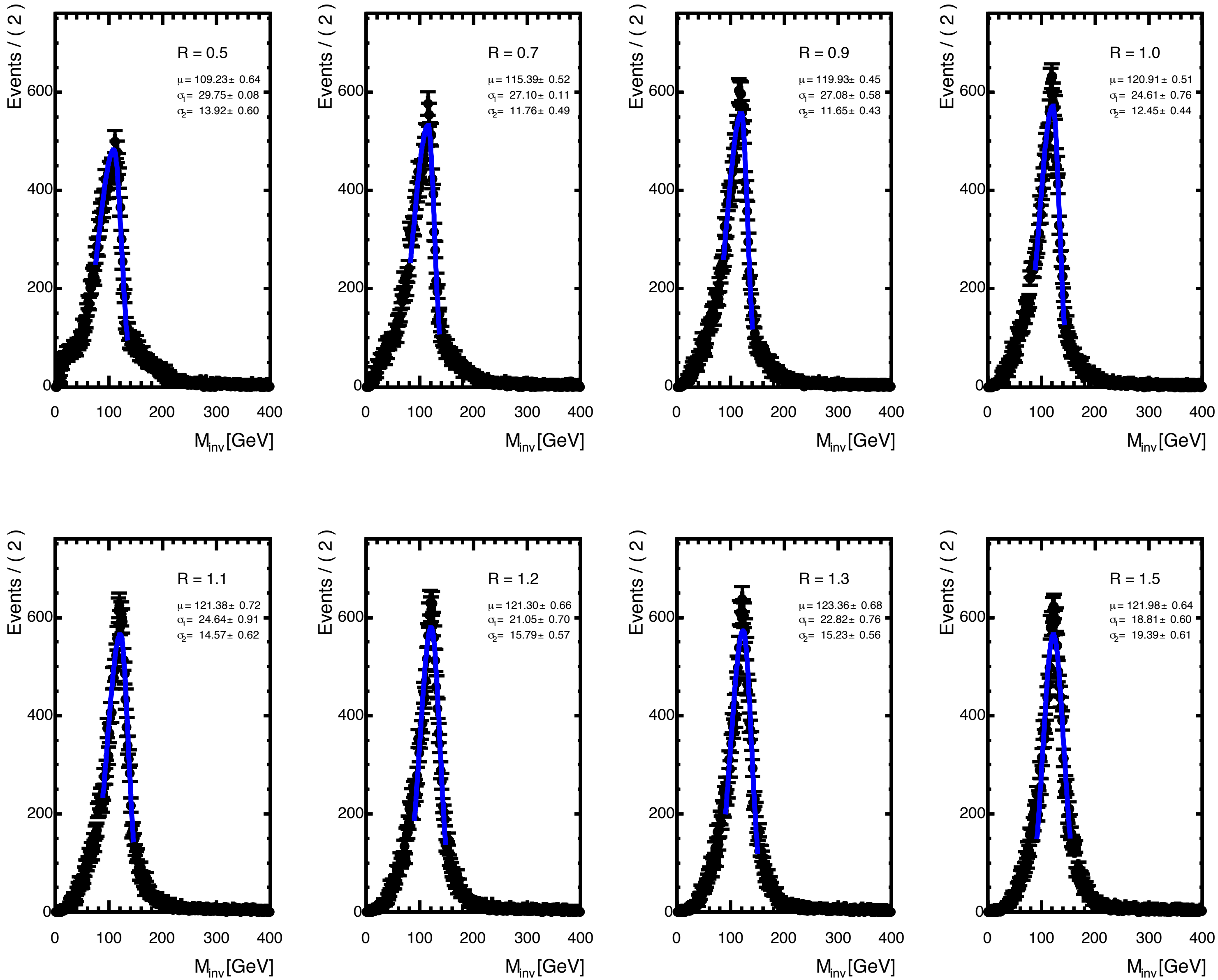
- **Software:** fastjet 3.1.2 + ValenciaPlugin contribution
- **Sample:** vector boson (W) fusion events hhnunu at 3 TeV, selected only events in which both Higgs decay in bb
- **Best chi2 mass:** mass computed minimising the chi2

$$\chi^2 = (m_{\alpha,\beta} - m_H)^2 + (m_{\gamma,\delta} - m_H)^2 \qquad m_H = 126\ GeV$$

- **Truth matched association** to isolate the contribution of combinatory background has been performed (similar results, not shown here)
- **Mass fit:** asymmetric gaussian → 1 mean, 2 sigmas

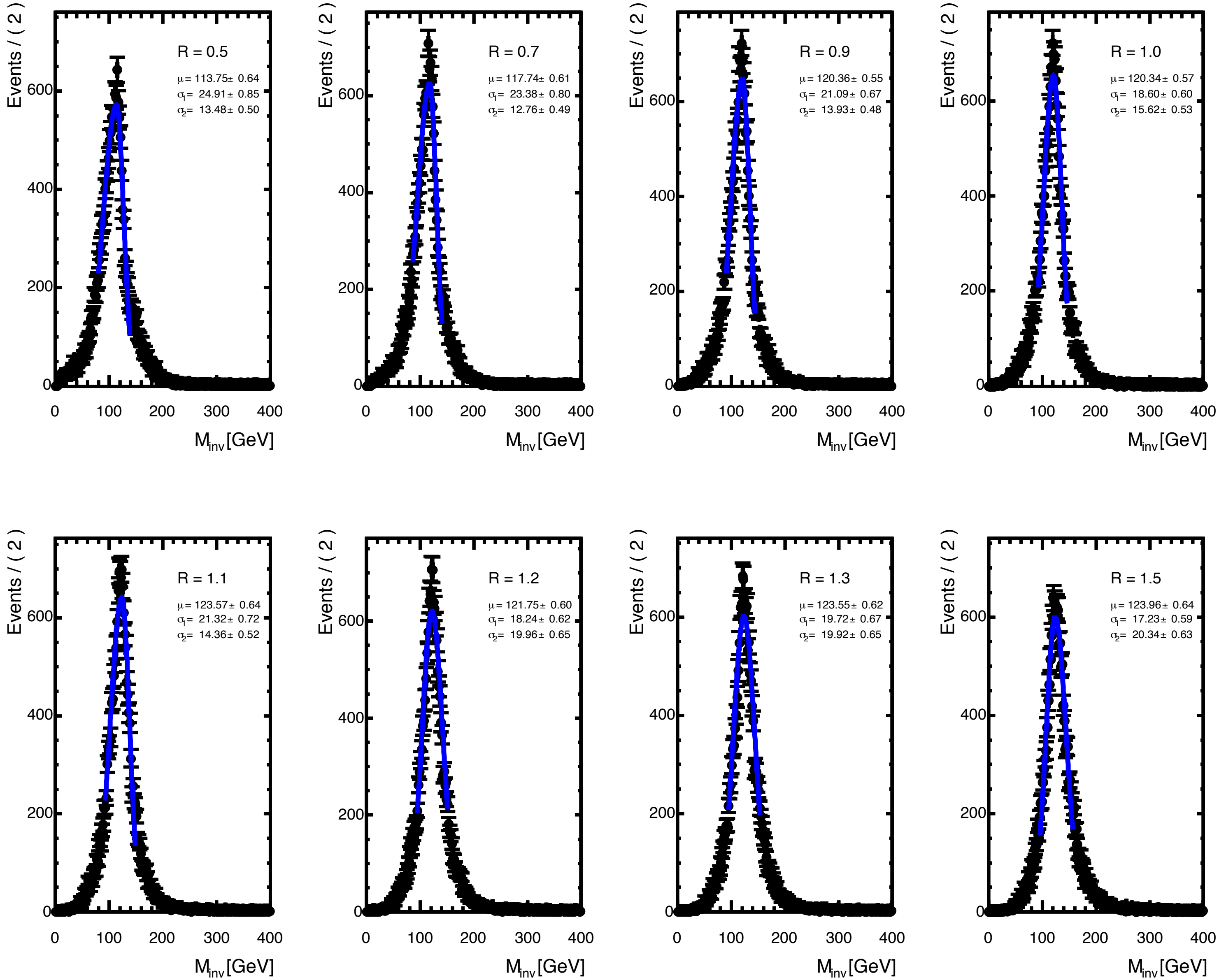
vlc – 4 jets

Asymmetric gaussian fit:
1 common mean, 2 sigmas



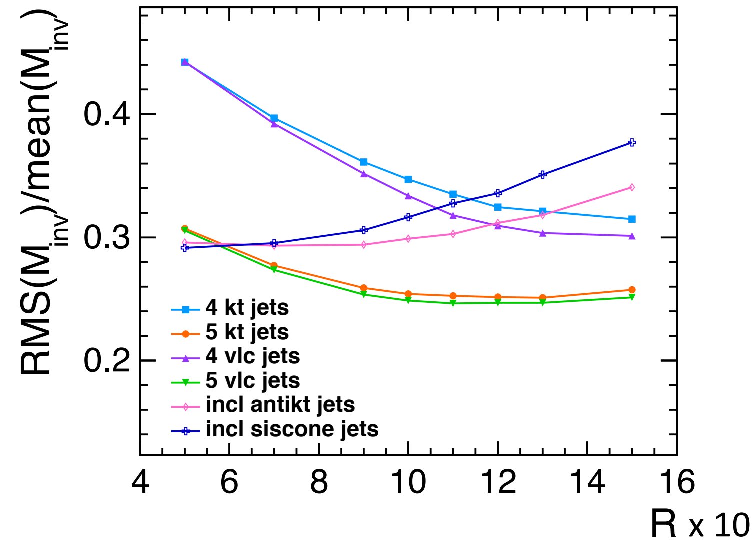
vlc – 5 jets

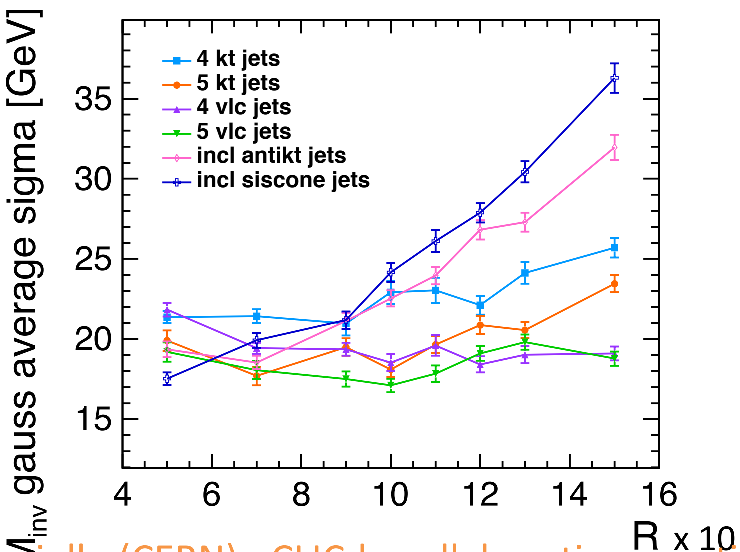
Asymmetric gaussian fit:
1 common mean, 2 sigmas

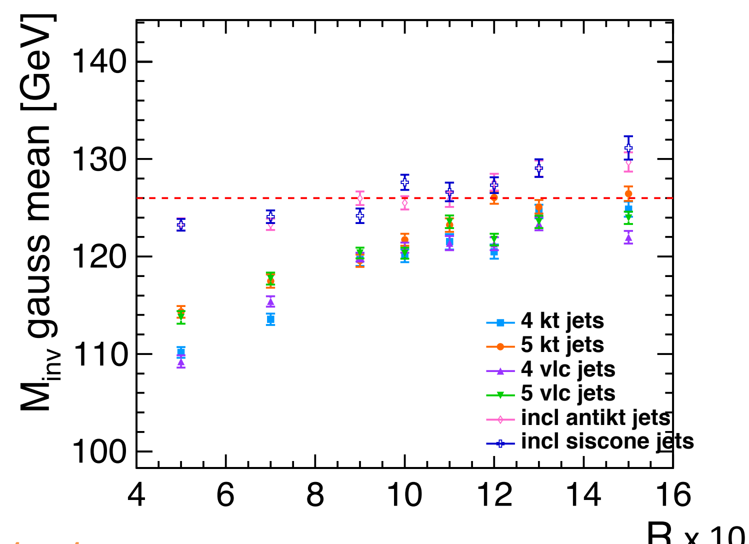


Results (χ^2 mass)

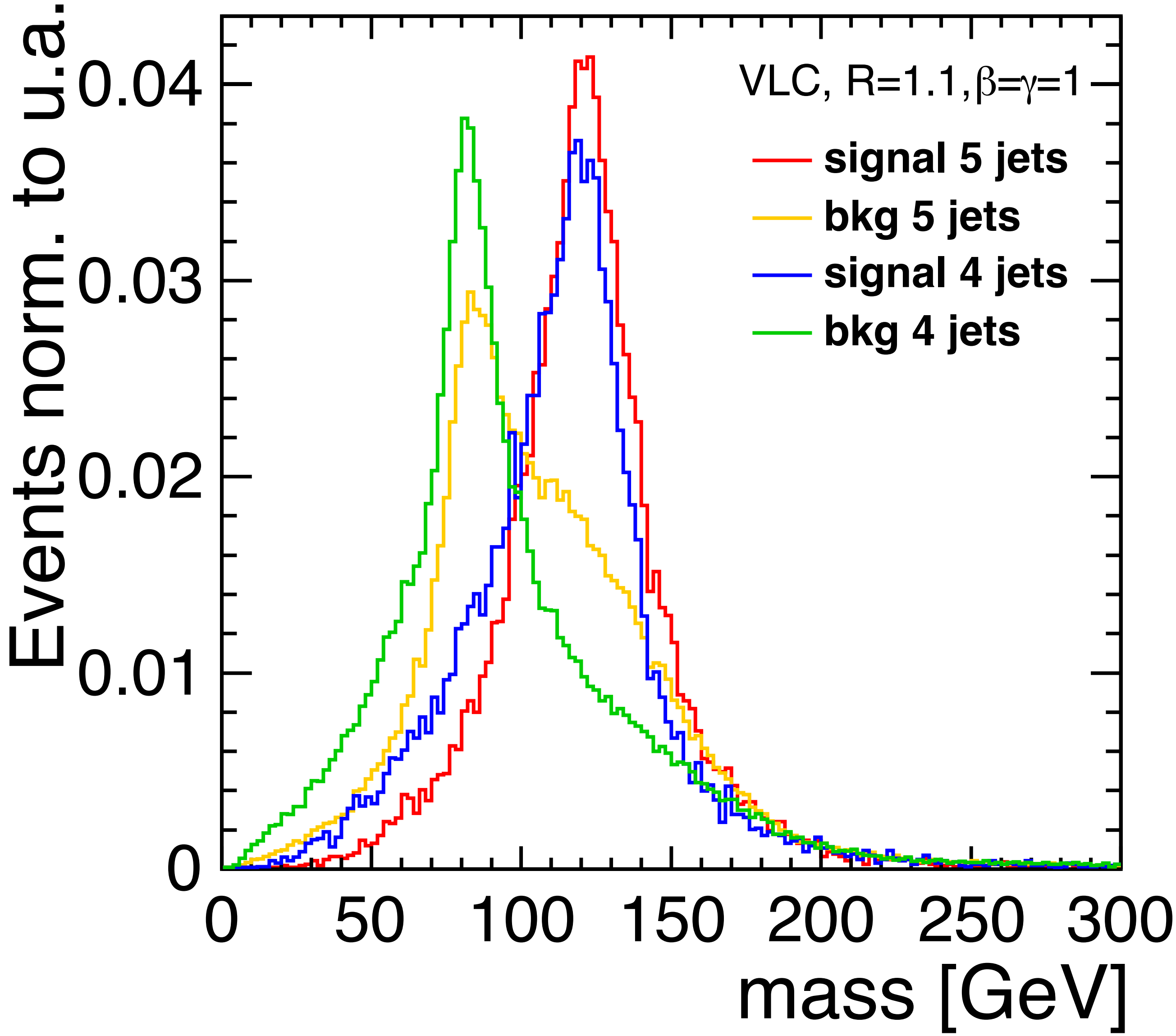
- Coefficient of variation: RMS/μ
  - 5 jets case improves up to 20-15% w.r.t. 4 jets
- Width of the mass distribution estimated with the average of the two fitted gaussian sigma.
- Fitted mean value of the mass a bit underestimated due to:
  - known bias in SiD calibration
  - presence of neutrinos in b jets







But bkg wants its part…



Still convenient going with 5 jets?

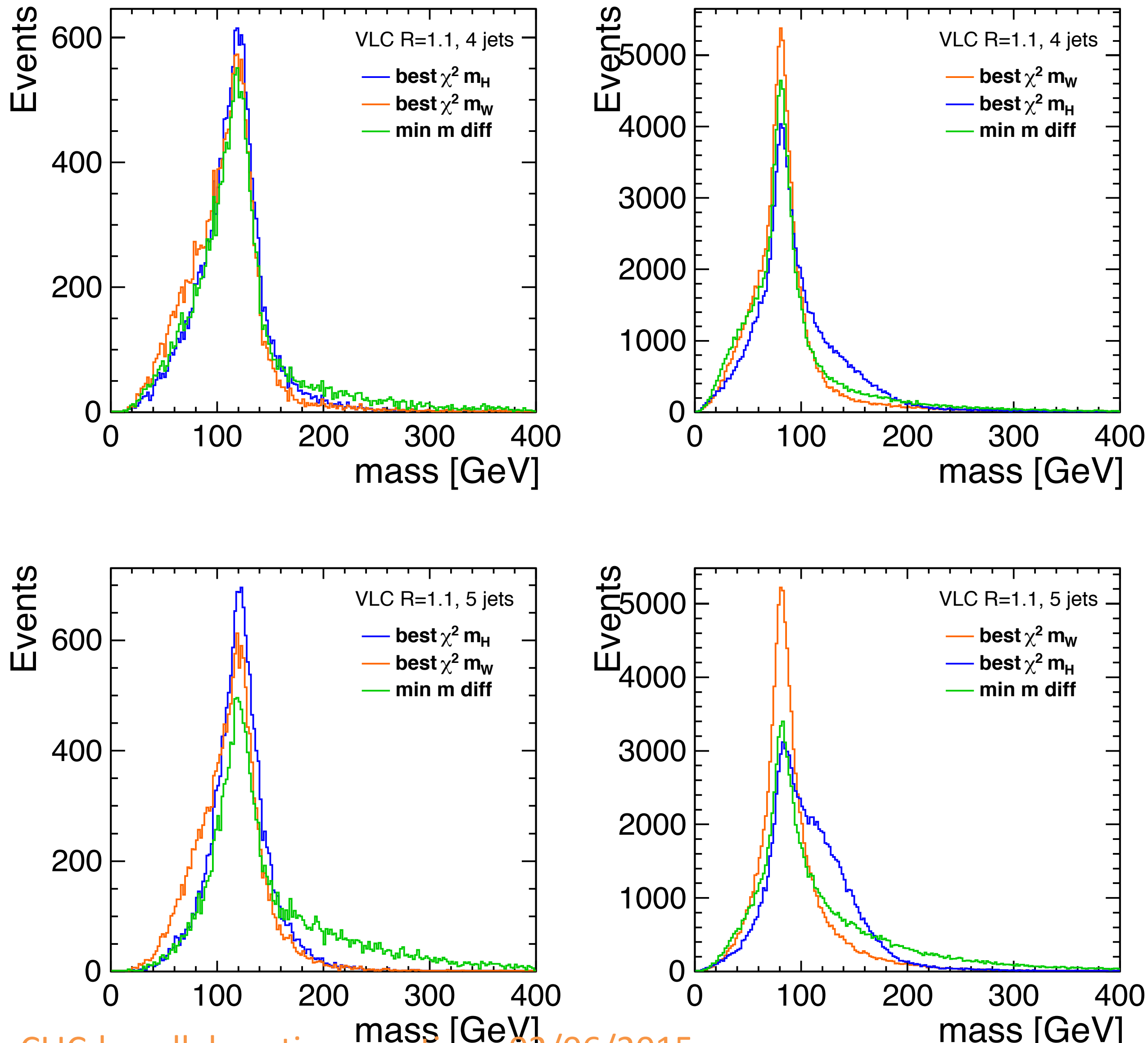
Can we be smarter than bkg?

min chi2 imposing mH
min chi2 imposing mW
min m difference

Higgs evts | W evts

4 jets →

5 jets →



bTag

- At 3 TeV, old slow approach used:
  - Run Jet reconstruction first,
    then vertexing,
    then Durham algorithm in LCFIPlus
- Doing vertexing as first step at 3 TeV would collect to much pile-up → worsening in mass resolution
- Done: training on Z→bb, Z→cc, Z→uu/dd/ss produced in WW fusion at 3 TeV
- Run Flavor Tag on signal and background samples → running on the grid now to get the final results
  - Run both on the 4 jets and 5 jets (in both cases 4 b-jets)

Conclusion

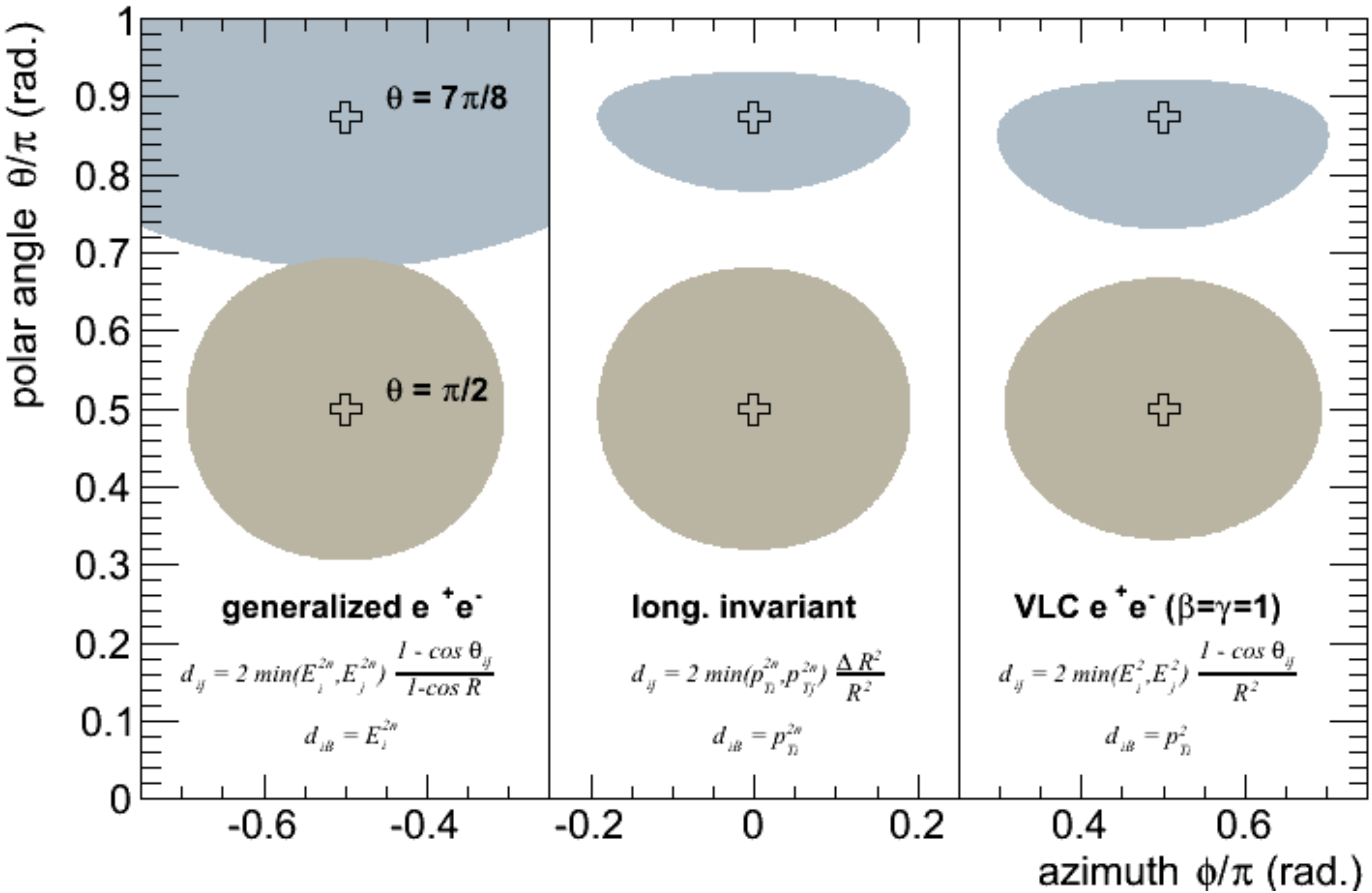
- Object optimisation at 3 TeV basically done
  - lot of effort invested in this phase to improve the analysis w.r.t. the past
- Lepton and Tau ID to maximise purity (used as veto in the analysis)
- Jet reconstruction algorithm:
  - VLC R=1.1 $\gamma=\beta=1$
  - Make final choice between 4 and 5 jets reclustering on the base of the best significance at the end of the analysis
- B-tag:
  - Vertexing after jet reconstruction (old approach)
  - Results running
- Ready to start MVA analysis → BDT

BACK-UP

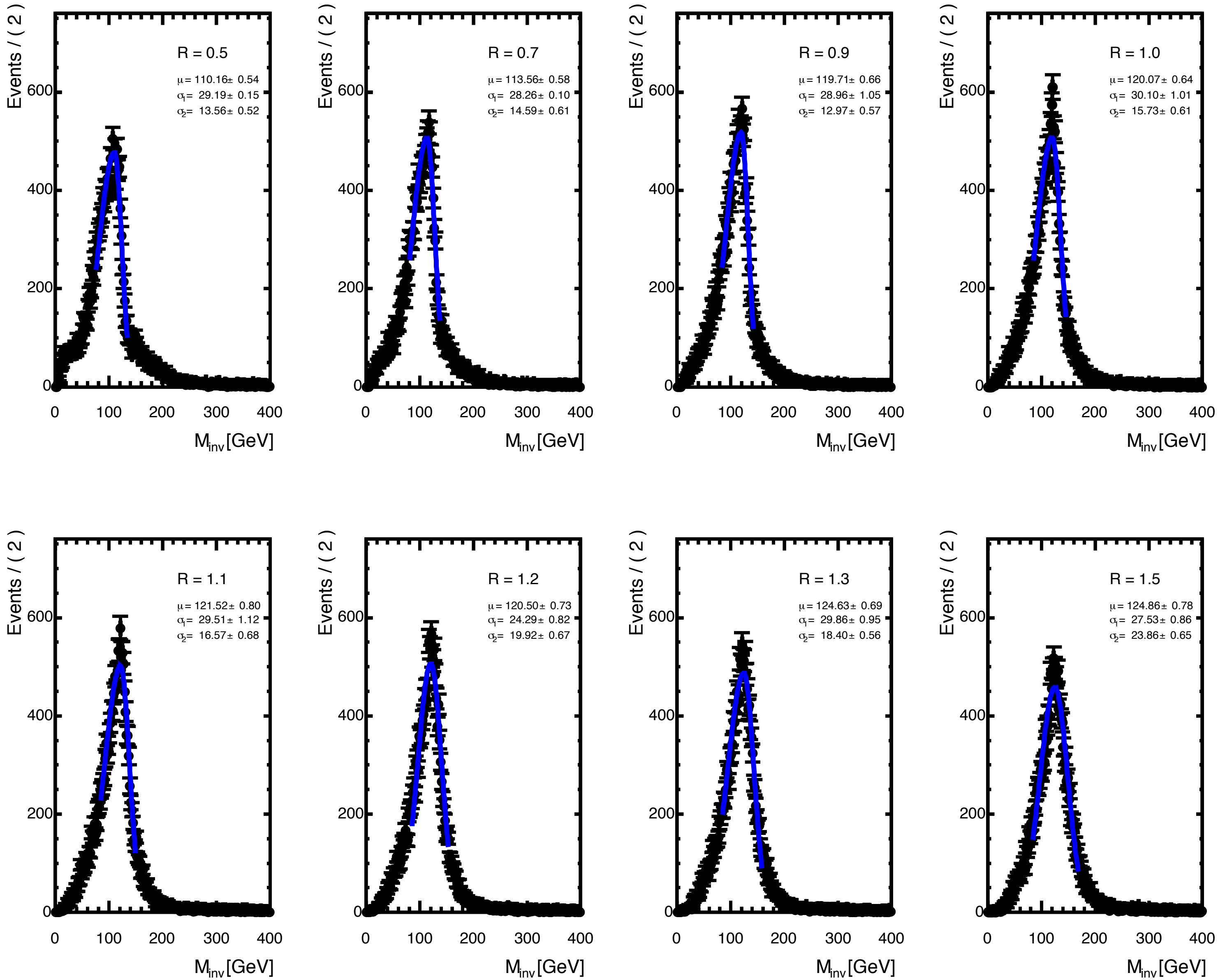
Truth match procedure

1. Start with reconstructed Jets
2. Ask for PFO particles
3. Link to MC particles
4. Follow the stable MC particle back to origin
5. If it come from b (from Higgs) marked as matched otherwise as not matched
6. Compute the fraction of matched particles in the jet over total number of particle in the jet

VLC – comparison of jet shapes

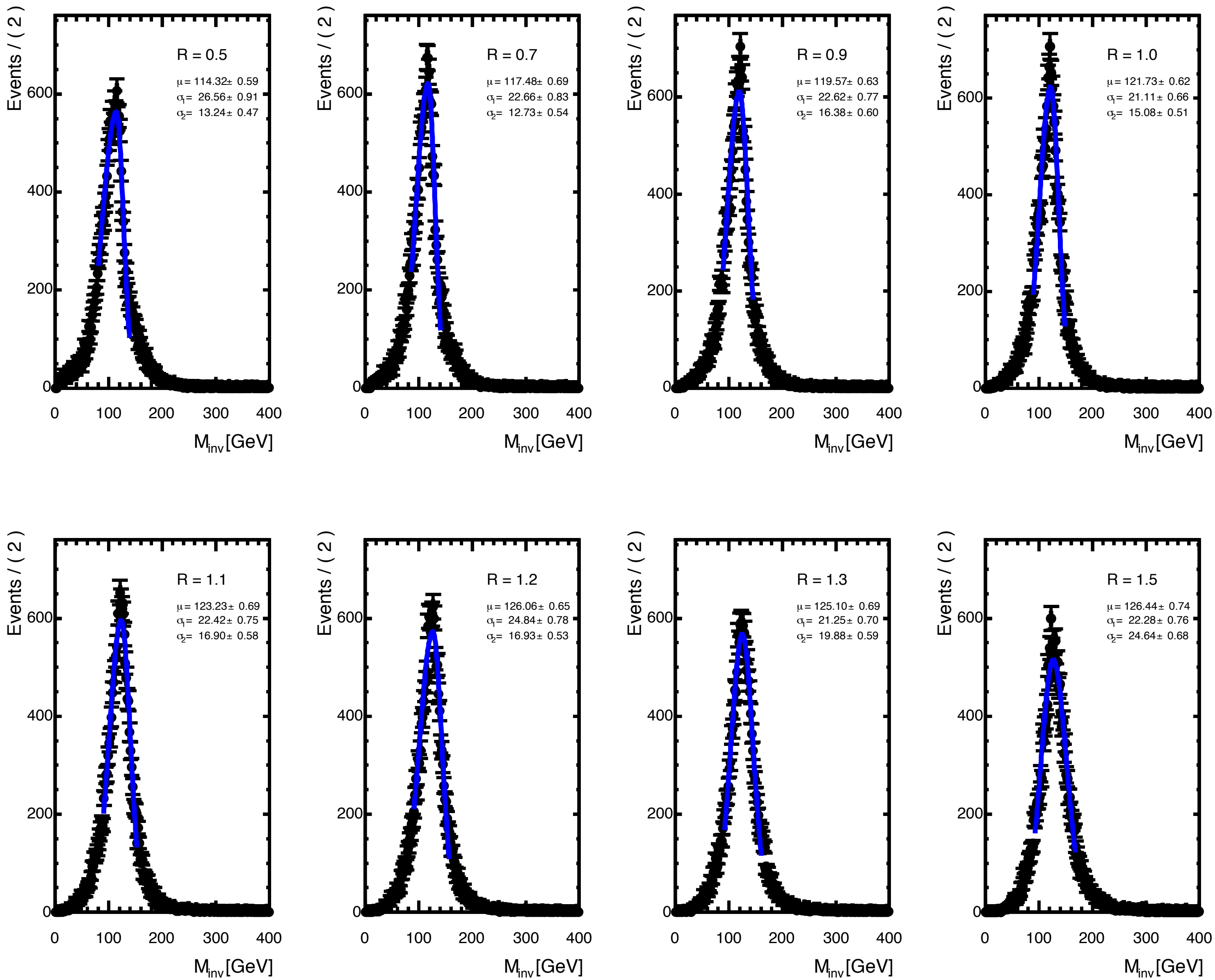


kt – 4 jets



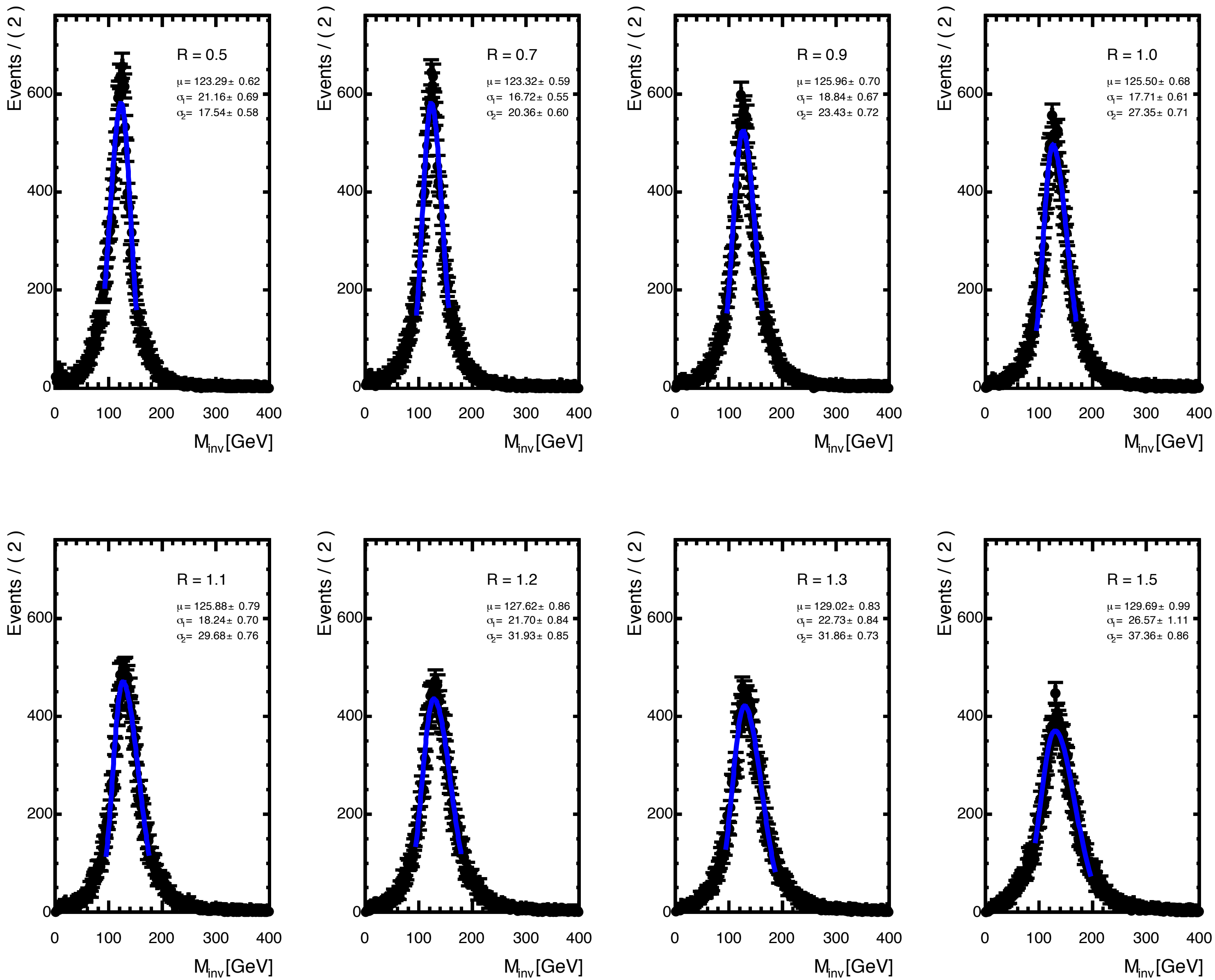
R = 0.5
μ = 110.16 ± 0.54
σ1 = 29.19 ± 0.15
σ2 = 13.56 ± 0.52
R = 0.7
μ = 113.56 ± 0.58
σ1 = 28.26 ± 0.10
σ2 = 14.59 ± 0.61
R = 0.9
μ = 119.71 ± 0.66
σ1 = 28.96 ± 1.05
σ2 = 12.97 ± 0.57
R = 1.0
μ = 120.07 ± 0.64
σ1 = 30.10 ± 1.01
σ2 = 15.73 ± 0.61
R = 1.1
μ = 121.52 ± 0.80
σ1 = 29.51 ± 1.12
σ2 = 16.57 ± 0.68
R = 1.2
μ = 120.50 ± 0.73
σ1 = 24.29 ± 0.82
σ2 = 19.92 ± 0.67
R = 1.3
μ = 124.63 ± 0.69
σ1 = 29.86 ± 0.95
σ2 = 18.40 ± 0.56
R = 1.5
μ = 124.86 ± 0.78
σ1 = 27.53 ± 0.86
σ2 = 23.86 ± 0.65
Events / (2)
Minv [GeV]
0
200
400
600
0
100
200
300
400

kt – 5 jets

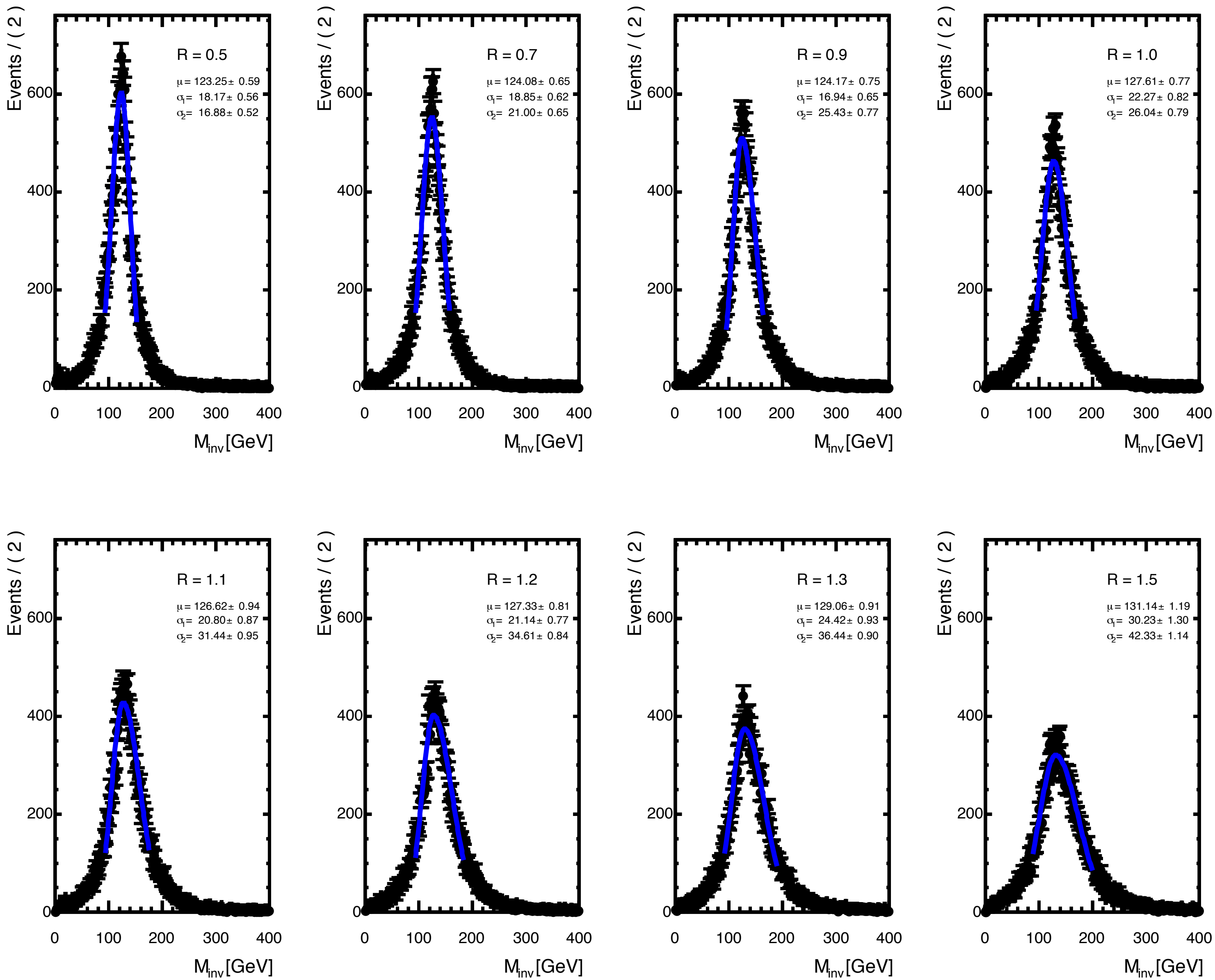

R = 0.5
μ = 114.32± 0.59
σ1 = 26.56± 0.91
σ2 = 13.24± 0.47
R = 0.7
μ = 117.48± 0.69
σ1 = 22.66± 0.83
σ2 = 12.73± 0.54
R = 0.9
μ = 119.57± 0.63
σ1 = 22.62± 0.77
σ2 = 16.38± 0.60
R = 1.0
μ = 121.73± 0.62
σ1 = 21.11± 0.66
σ2 = 15.08± 0.51
R = 1.1
μ = 123.23± 0.69
σ1 = 22.42± 0.75
σ2 = 16.90± 0.58
R = 1.2
μ = 126.06± 0.65
σ1 = 24.84± 0.78
σ2 = 16.93± 0.53
R = 1.3
μ = 125.10± 0.69
σ1 = 21.25± 0.70
σ2 = 19.88± 0.59
R = 1.5
μ = 126.44± 0.74
σ1 = 22.28± 0.76
σ2 = 24.64± 0.68
Events / (2)
M_inv [GeV]
0
100
200
300
400
600

Asymmetric gaussian fit:
1 common mean, 2 sigmas

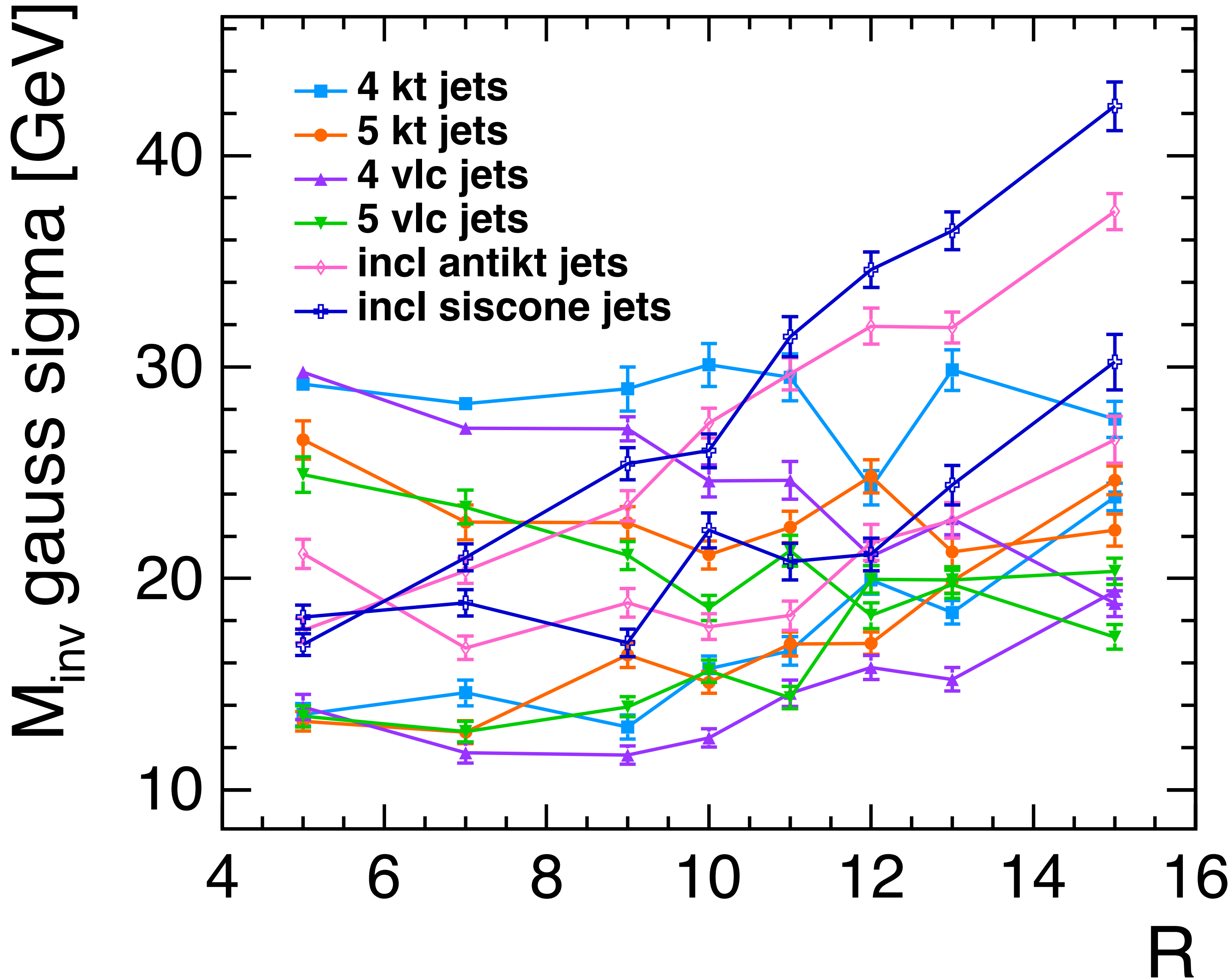
antikt – p_T^{min} = 4 GeV



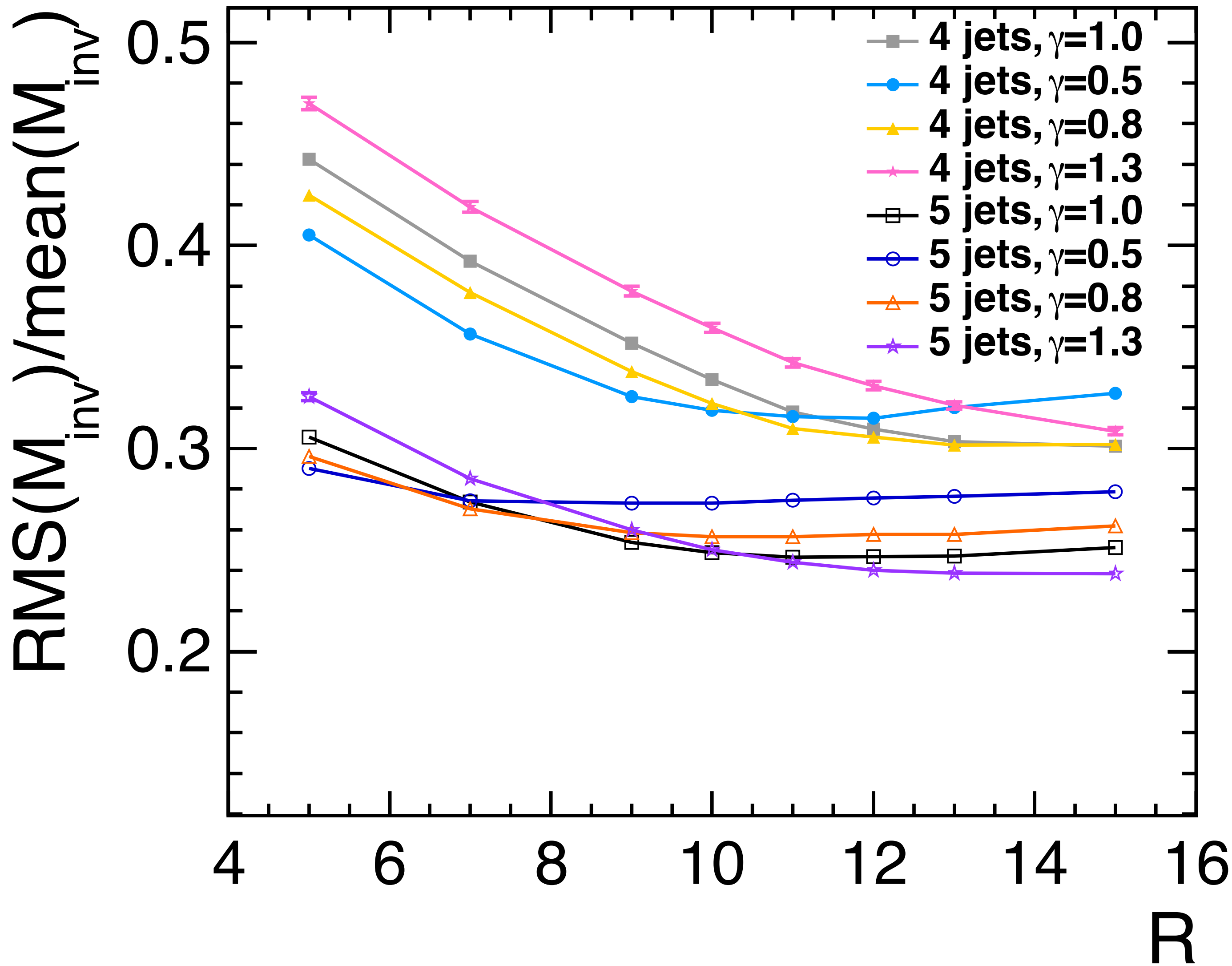
siscone - p_T^{min} = 4 GeV



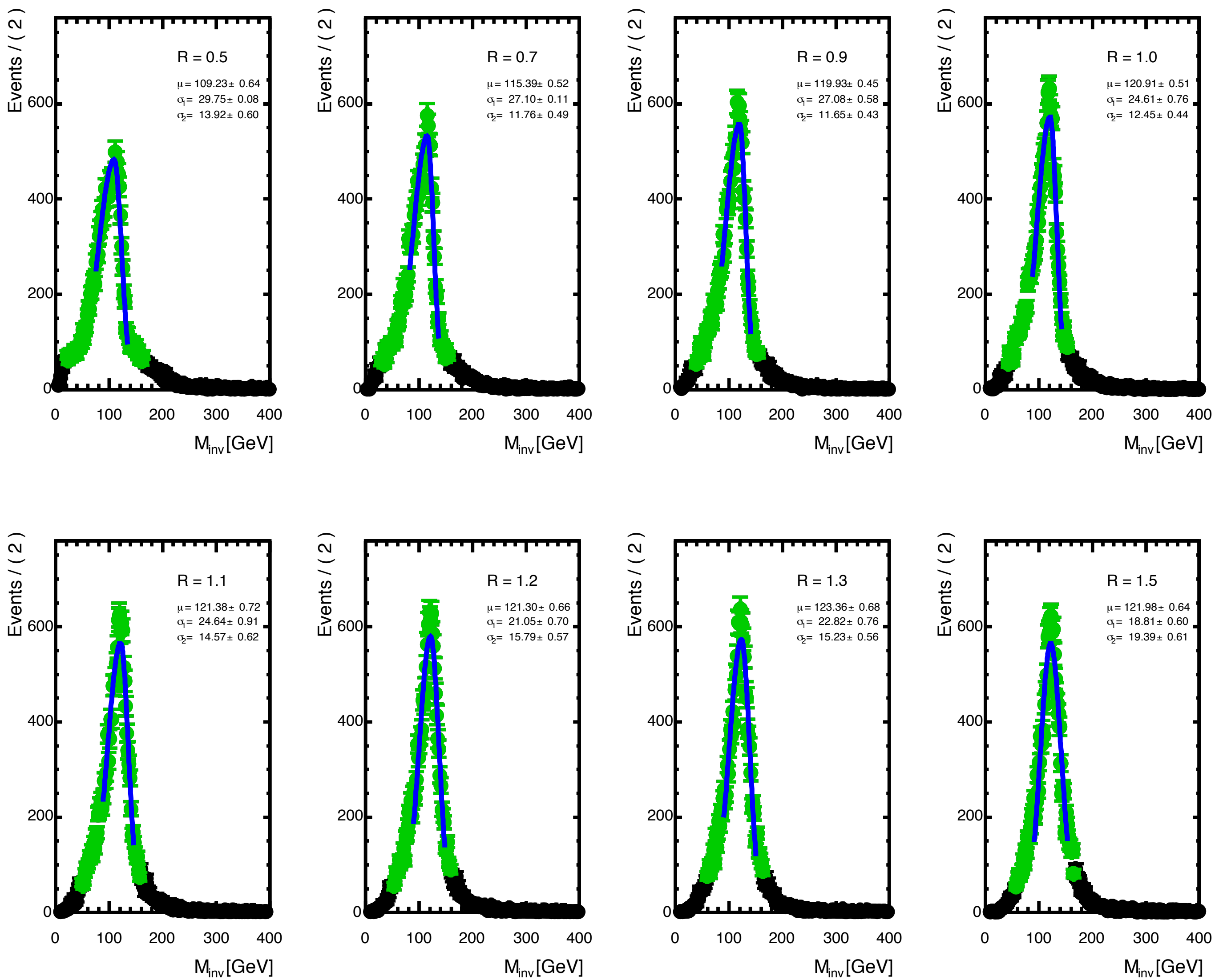
Left and right gauss resolution



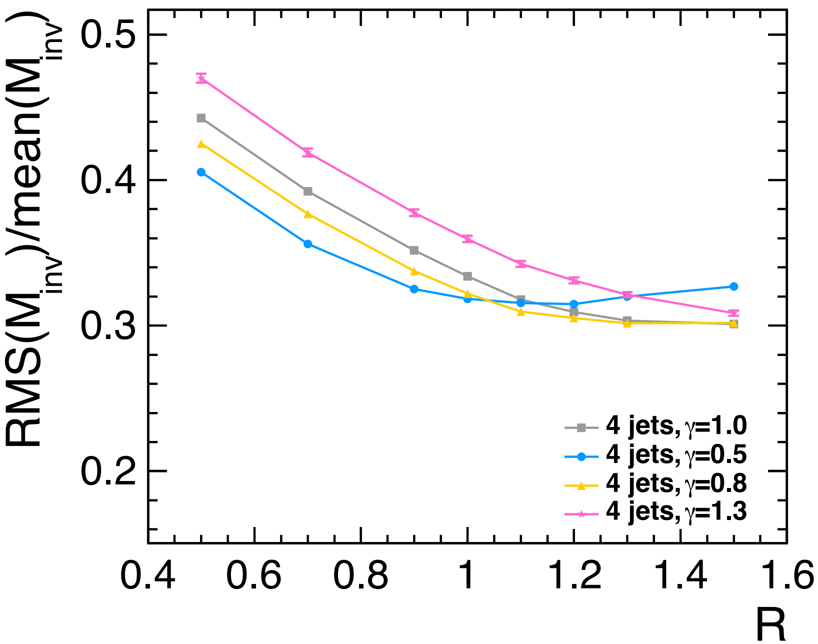
VLC – gamma comparison

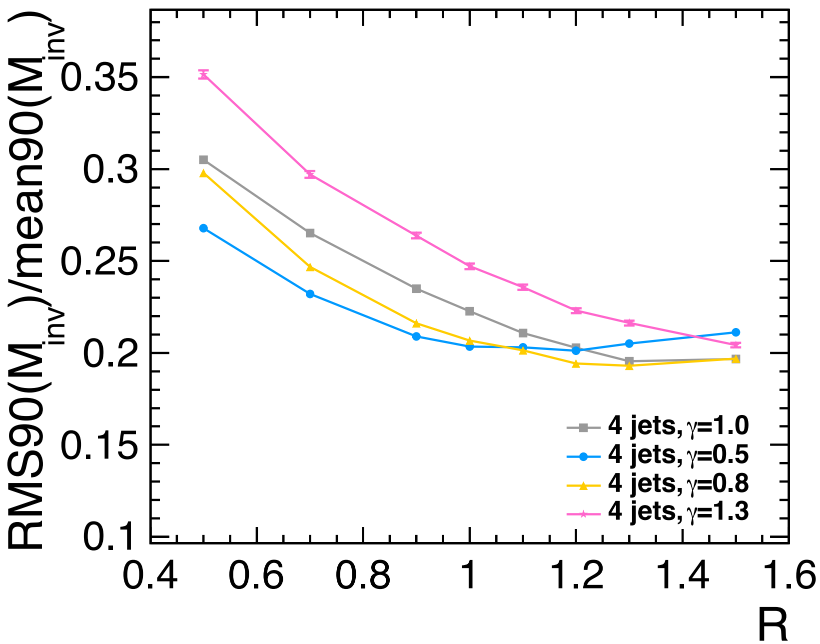


mass distributions for VLC algorithm



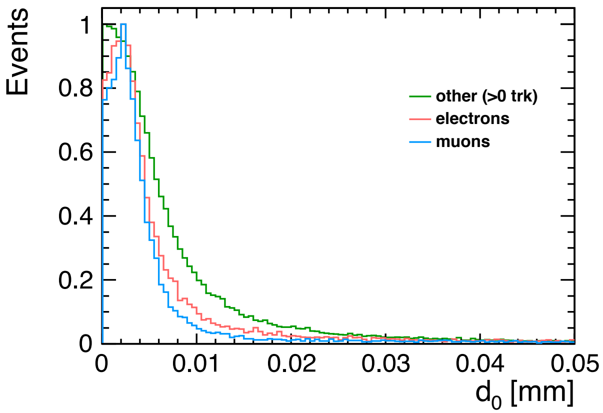
coefficient of variation

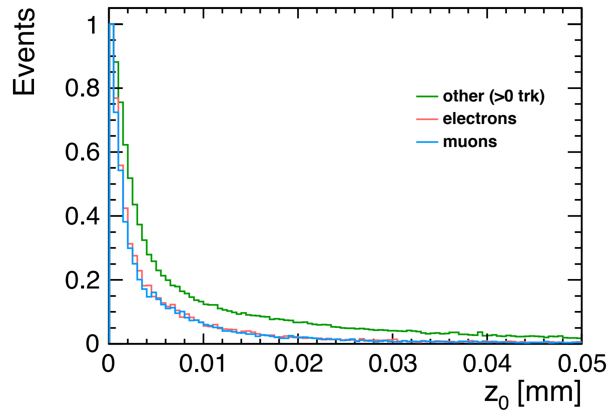


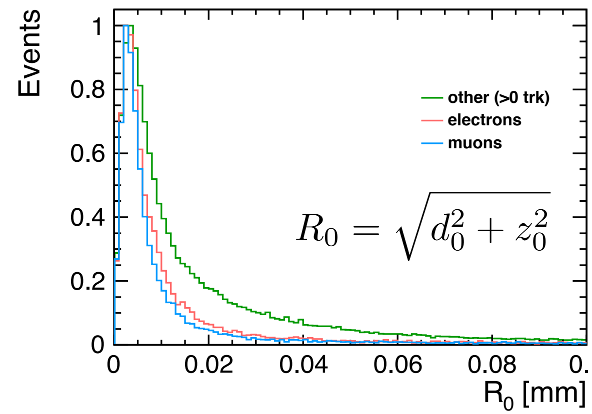


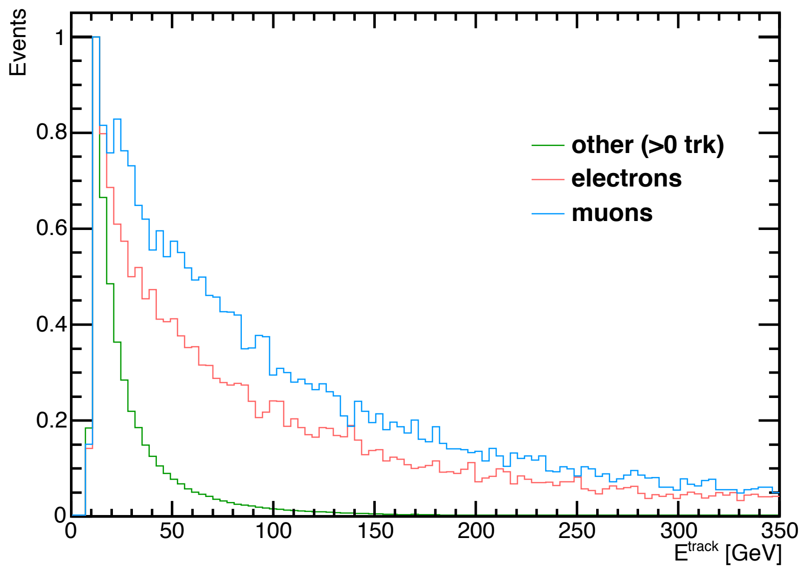
Lep ID – basic variables

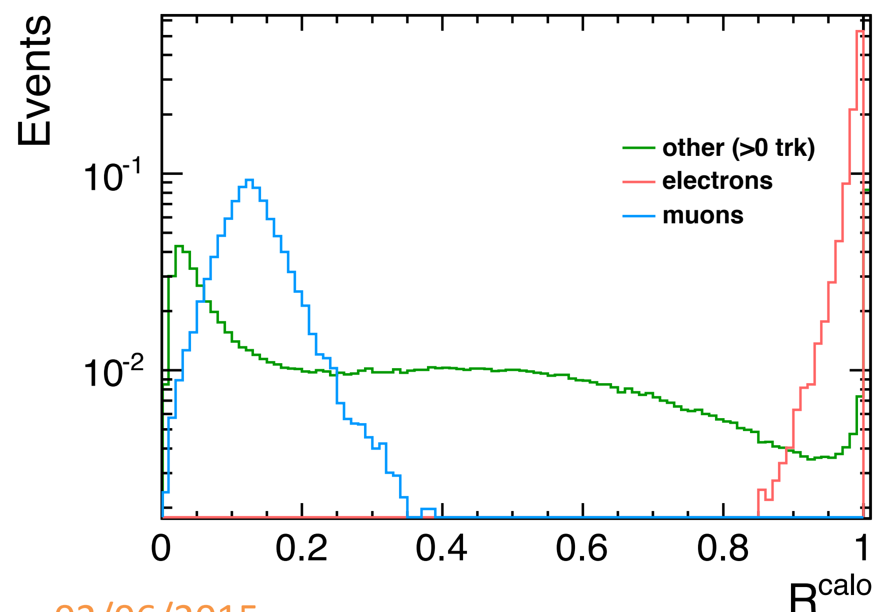
- d0<0.03, z0<0.04, R0<0.06
- 0.05<Rcal<0.25 || Rcal>0.9



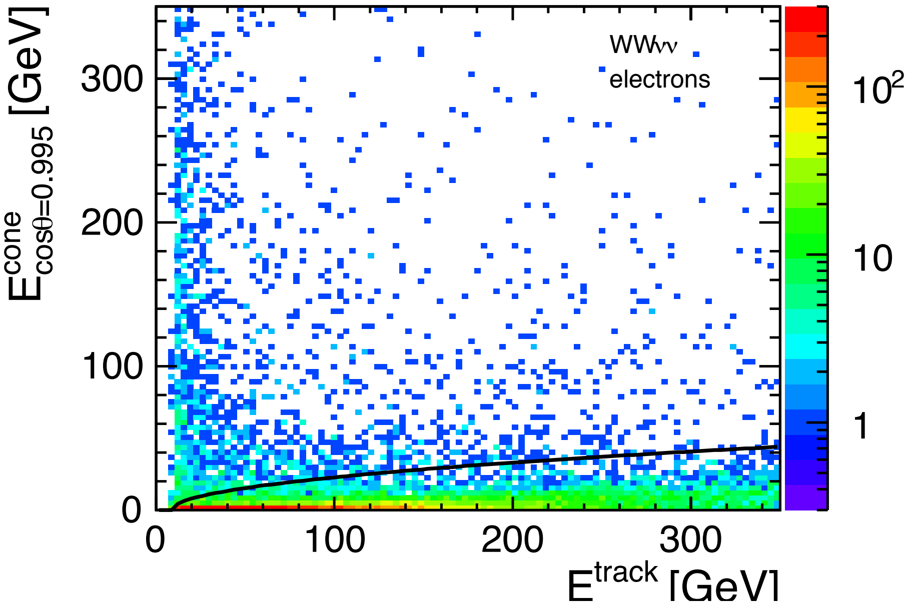


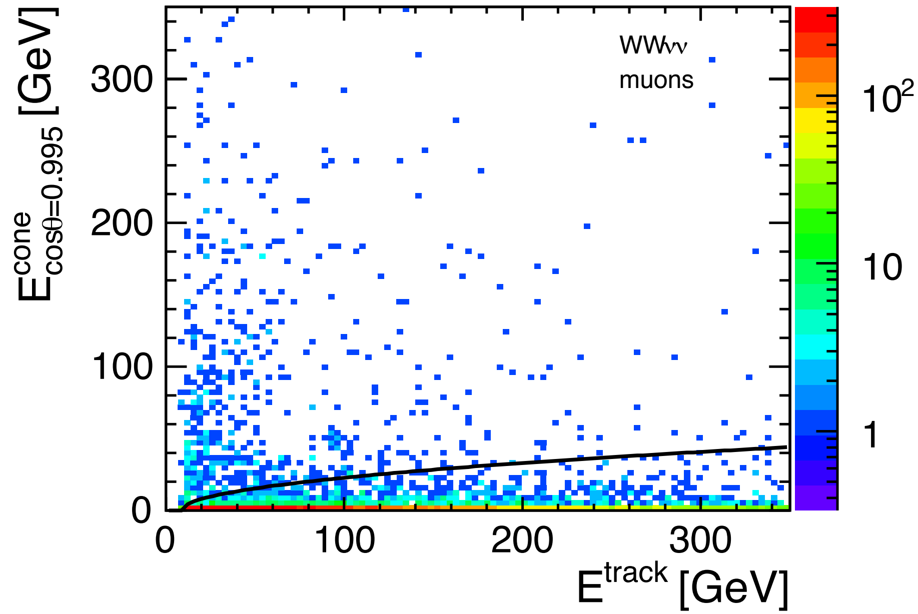


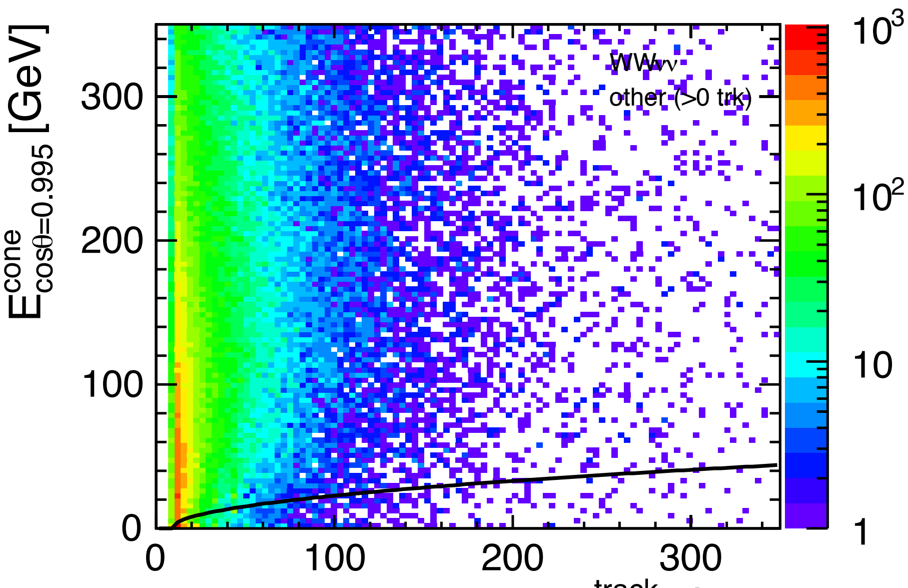




Lep ID – polynomial isolation







In these plots cuts on variable on previous slide are already applied

A=0.0, B=5.7, C=-50.

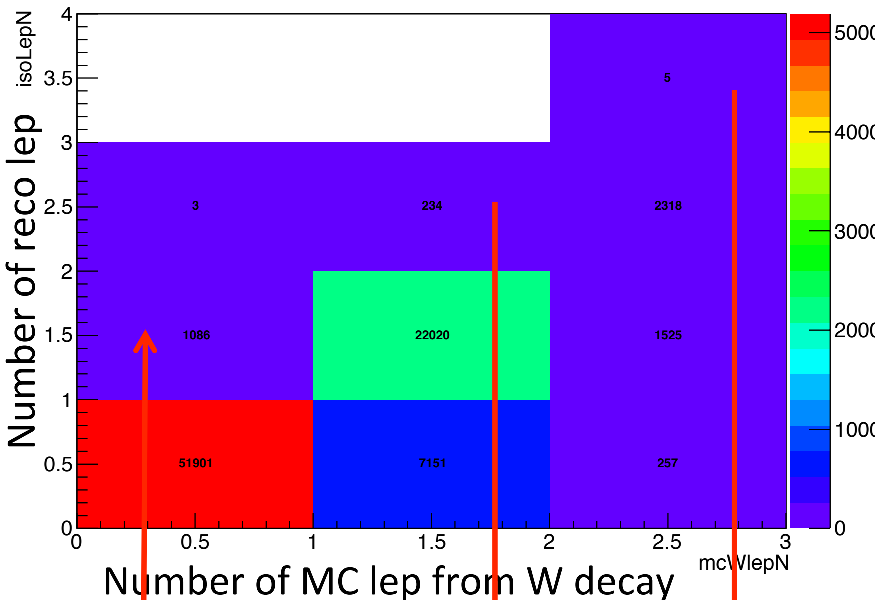
$$E_{cone}^2 <= A*E_{track}^2 + B*E_{track} + C$$

energies are in GeV

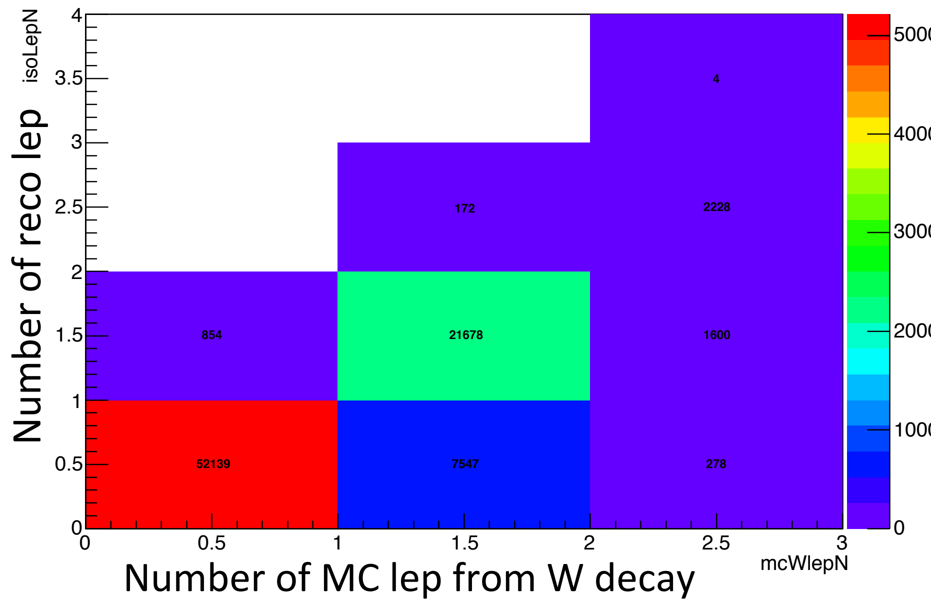
Lep ID – test on WWnunu sample

- In the analysis: veto of events with at least one iso lep to reject WWnunu
- *Problem in Pandora: an electron may be splited up to 4 contributions (usually 2) → rejected the ones without track*

cut on Etrack = 10 GeV



Extra cut on Etrack = 15 GeV



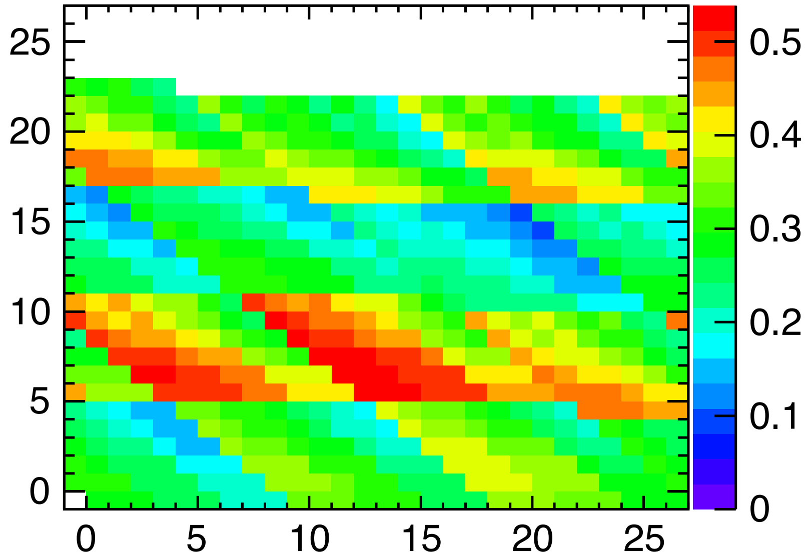
Not important in more than the expected leptons are reco in these bins, the events should be vetoed anyway.

Instead try to minimize fake lep reco in 0 lep events (final aim: avoid rejecting signal) → add extra cut on Etrack = 15 GeV

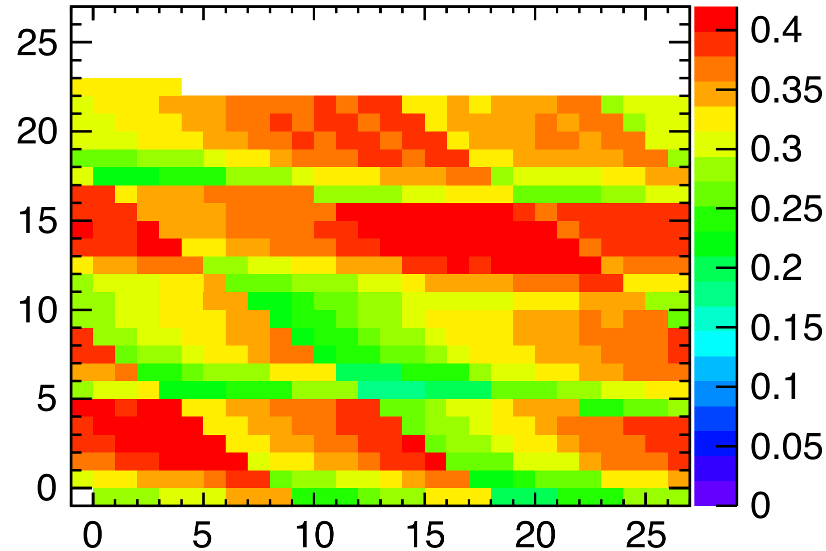
Tau ID

***ATT: these are NOT plots**, the axes have no meaning. They are just "tables". Each "bin" represents a tau collection reconstructed with different initial parameters. In total 648 tau different identification → first 25 plotted on the first row and so on…*

PURITY: n_matched_tau/n_reco_tau



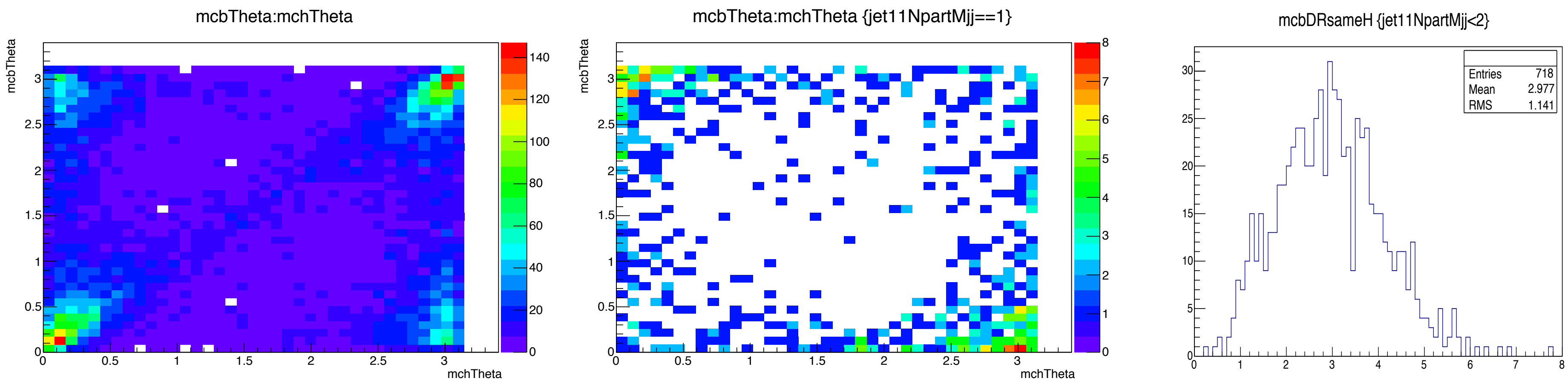
EFF: n_matched_tau/n_mc_tau



- Tau identification is used as veto → choose working point that maximize purity (and only subsequentially efficiency) to not loose signal efficiency

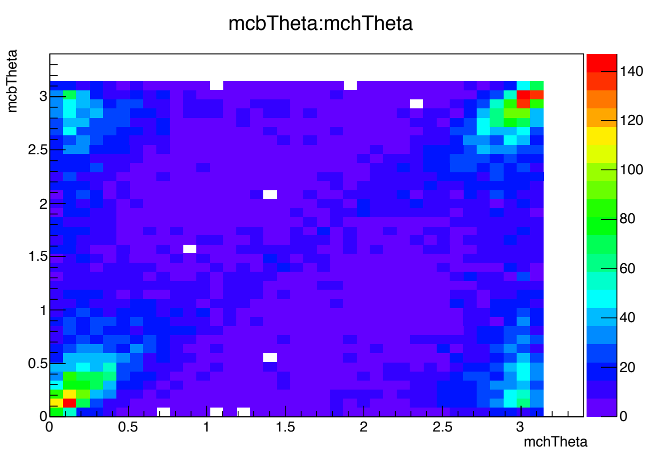
Comments: only 1 truth matched jet

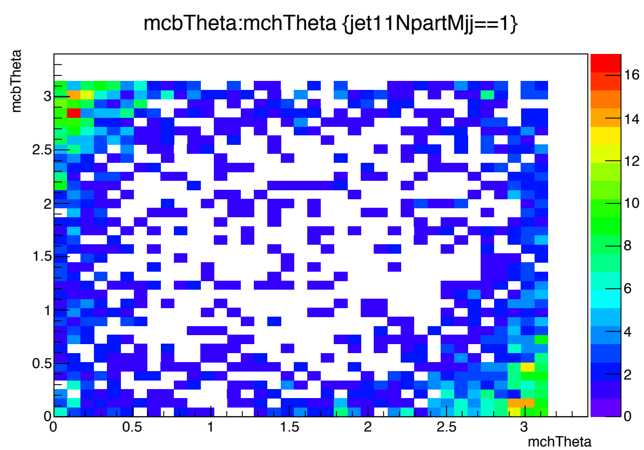
- Events where only 1 jet is truth matched to one of the Higgs:
  - Forward bs
  - Higgs and b in opposite direction → the Higgs decays mostly at rest → difficult to reconstruct
- Some conclusion from distribution of the DR between the bs coming form the same Higgs → the 2 bs are mostly back-to-back
- Plots for exclusive **5 VLC jets**

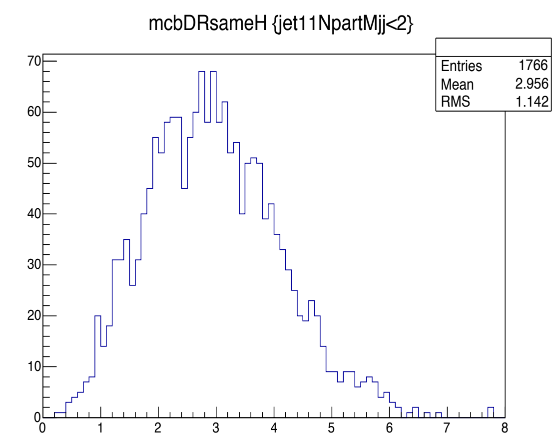


Comments: only 1 truth matched jet

- Events where only 1 jet is truth matched to one of the Higgs:
  - Forward bs
  - Higgs and b in opposite direction → the Higgs decays mostly at rest → difficult to reconstruct
- Some conclusion from distribution of the DR between the bs coming form the same Higgs → the 2 bs are mostly back-to-back
- Plots for exclusive **4 VLC jets**







Comments: average mass

- No strong correlation between the mass of the two Higgs in the same event (left plot kt 4 jets, right plot kt 5 jets)
- In this analysis there is not a large problem of contribution sharing between different jets, but of loosing part of the jets