

Measurement of hadronic Higgs boson branching ratios at FCC-ee with ZH events at $\sqrt{s}=240$ GeV

Giovanni Marchiori (APC Paris)

Work down with Alexis Maloizel and Paul Guimbard (ENS Paris-Saclay)

FCC France-Italie workshop

22 November 2022



Introduction

- **Goals:**

- estimate sensitivity of FCC-ee to hadronic branching ratios of Higgs boson (bb / cc / gg) (\Rightarrow couplings to b/c/g)
- compare sensitivity when different assumptions on efficiency/rejection of taggers (or alternative working points) are used

- **Analysis main features:**

- Started with $Z(\ell)H$, $\ell=e, \mu$ channel at $\sqrt{s}=240$ GeV first
 - Will look at other Z decay channels / \sqrt{s} later
- Assume an integrated luminosity of 5/ab
- Only statistical uncertainties are considered
- Reconstruction/selection performed on centrally produced FCCee MC samples (spring2021 production) using FCCAnalysis software; final fit performed with private code based on ROOT/RooFit

Current analysis status - more details

- Analysis relies on march2022 version of FCCAnalysis for event reconstruction (jet clustering, tagging ...)
 - With some ad-hoc modifications for instance to improve the efficiency of jet truth-flavour labelling
- Analysis was also re-run on privately generated samples to check the impact of using non-isolated leptons
 - When performing the analysis, we noticed a 30% efficiency drop due to the fact that the “Electrons” and “Muons” collections in the spring2021 production store leptons after an isolation requirement
 - Privately generated samples were produced with the same generator and Pythia cards as the official samples, just saving in output also lepton collections (AllElectrons, AllMuons) without the isolation requirement
- Migration of analysis code to latest FCCAnalysis version and analysis model/scripts (stage1, stage2, ..) recently performed, but analysis not rerun yet (plan is to run directly with improvements such as latest developments on flavour tagging)

Analysis strategy

- The measurements proceeds in the following steps:
 - Event **reconstruction**: leptons, jets and missing energy are reconstructed
 - Event **selection**: events consistent with the signature under study are kept
 - NOTE: no systematic optimisation of the selection criteria has been performed
 - Event **categorisation**: selected events are classified in categories based on # of b-, c- and g-tagged jets
 - **Fit** for BR measurement:
 - the signal yield in each category and for each Higgs decay mode (bb, cc, gg, non-hadronic) is extracted through a **simultaneous extended maximum likelihood fit to the recoil mass distribution of the various tagging categories**
 - Assuming tagging efficiencies for each flavour type to be known, the acceptance of each category for the various Z(II)H(XX) is known and the system of equations relating the yields to the product σBR can be solved
 - In practice, in the likelihood the yields are expressed directly in terms of products of σBR times the acceptances and the fit returns σBR

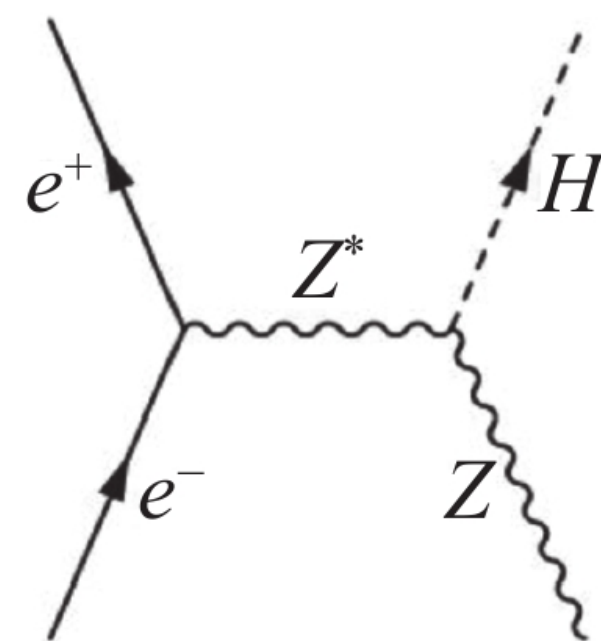
Samples

- Use samples from central production ([link](#))
 - **ZZ**: Pythia8, 60M
 - **WW**: Pythia8, 10M
 - **Z/ γ^* (ee) [Mee in 30-150 GeV]**: Whizard + Pythia6, 79M
 - **Z/ γ^* ($\mu\mu$)**: Whizard + Pythia6, 49M
 - **Z(l)H(bb/cc/gg/others)**: Whizard + Pythia6, 1M $\mu\mu$ H + 0.9M eeH
- To have more of H(cc) and H(gg), we produced (with the same settings) samples separated by Higgs decay:
 - llH(bb): 100k llH(cc): 100k llH(gg): 100k llH(other): 500k
- ISR, FSR, beam energy spread are turned on
- The simulated detector is the FCC-ee IDEA one (spring2021 Delphes card)

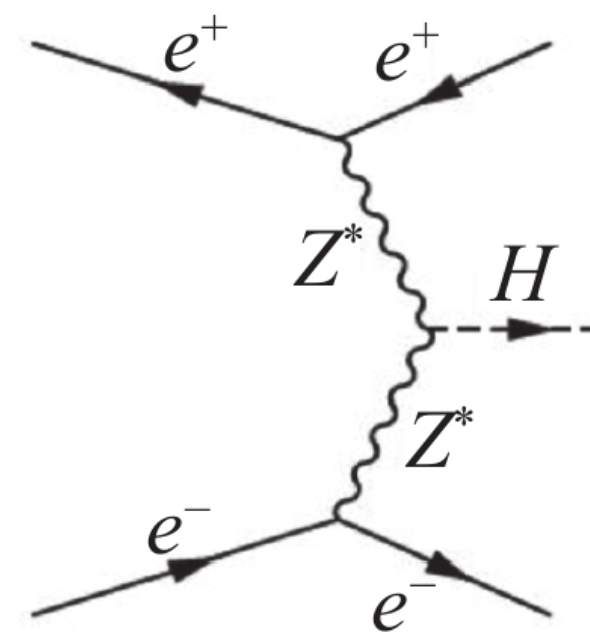
Samples (cont'd)

| Process | sigma (fb) | BR1 | BR2 | sigma*BR (fb) | Ngen | LumiGen (fb-1) | LumiGen/Lumi |
|---------------|------------|----------|---------|---------------|----------|----------------|--------------|
| Z(l)H(bb) | 201.87 | 0.067316 | 0.5824 | 7.914280728 | 100000 | 12635 | 2.527 |
| Z(l)H(cc) | 201.87 | 0.067316 | 0.02891 | 0.3928603294 | 100000 | 254543 | 50.909 |
| Z(l)H(gg) | 201.87 | 0.067316 | 0.08187 | 1.112538055 | 100000 | 89885 | 17.977 |
| Z(l)H(nonhad) | 201.87 | 0.067316 | 0.30682 | 4.169401808 | 500000 | 119921 | 23.984 |
| ZZ | 1358.99 | 1 | 1 | 1358.99 | 59800000 | 44003.26713 | 8.801 |
| WW | 16438.5 | 1 | 1 | 16438.5 | 10000000 | 608.3280105 | 0.122 |
| Z/gamma(ee) | 8305 | 1 | 1 | 8305 | 79400000 | 9560.505719 | 1.912 |
| Z/gamma(mm) | 5288 | 1 | 1 | 5288 | 49400000 | 9341.906203 | 1.868 |

- Signal xsections are corrected to take into account that whizard samples also include small contribution from ZZ fusion



6.76 fb

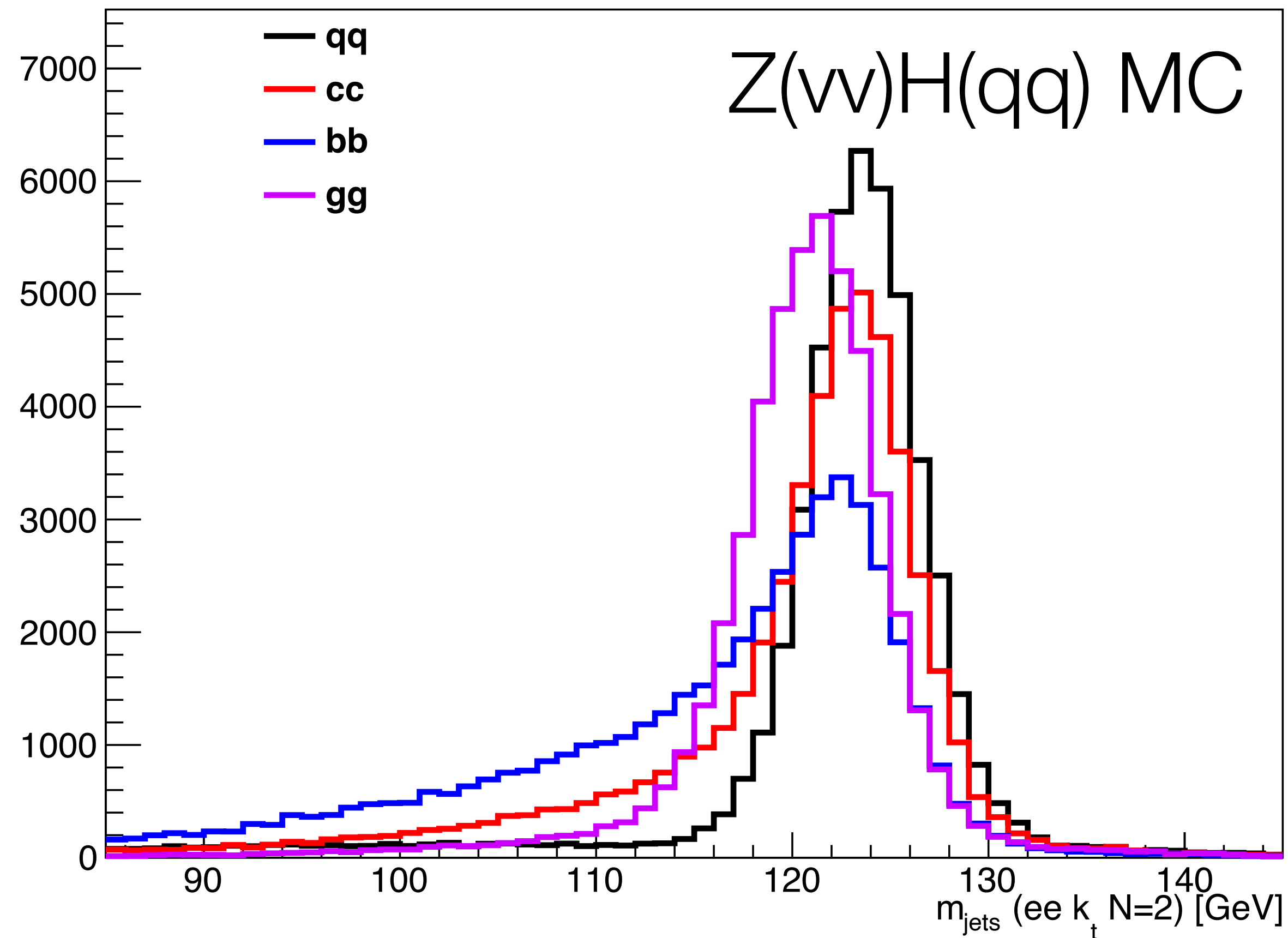


0.40 fb

| | sigma*BR (fb) | Ngen | LumiGen (fb-1) | LumiGen/Lumi |
|-------------|---------------|--------|----------------|--------------|
| llH(bb) | 8.150082527 | 100000 | 12270 | 2.454 |
| llH(cc) | 0.4045653947 | 100000 | 247179 | 49.436 |
| llH(gg) | 1.145685537 | 100000 | 87284 | 17.457 |
| llH(nonhad) | 4.293626925 | 500000 | 116452 | 23.290 |

Jet reconstruction

- For jet reconstruction, use **ee exclusive k_t algorithm with $N=2$, E-recombination scheme**
- Leads to good invariant mass resolution for Higgs hadronic decays (modulo for heavy flavour decays decaying to neutrinos). Small shift in $H(gg) \Rightarrow$ imperfect clustering of particles in the shower?



Jet flavour labelling

- For jet flavour labelling, have noticed that standard algorithm in FCCAnalysis was not super efficient for all Higgs channels, so investigated alternative strategies, including recently proposed ghost matching
 - 1: default algorithm in FCC analysis (angular matching to partons, $\Delta\theta=0.3$, favour $b>c>l>g$)
 - 2: angular matching to partons, $\Delta\theta=0.3$, highest-E parton
 - 3: angular matching to partons, $\Delta\theta=0.3$, closest parton
 - 4: angular matching to partons, $\Delta\theta=0.3$, lowest $\Delta\theta/E$
 - 5: ghost matching to partons
 - 6: ghost matching to hadrons
 - 7: like 2 but use $\Delta\theta=0.8$
- Evaluated on $Z(\nu\nu)H(xx)$ events
 - efficiency defined as number of jets with the correct label over $2 \cdot N_{\text{events}}$, for instance:
 - $H(bb)$: 50k evts \rightarrow 100k b-jets \rightarrow eff = $N(\text{b-labelled jets})/100k$
- Decided to **use strategy 7: angular matching to partons, $\Delta\theta=0.8$, highest-E parton**

| Strategy | ϵ_b | ϵ_c | ϵ_l | ϵ_g |
|----------|--------------|--------------|--------------|--------------|
| 1 | 95.9 | 94.4 | 94.6 | 75.0 |
| 2 | 95.8 | 94.4 | 95.0 | 92.1 |
| 3 | 93.5 | 89.2 | 89.3 | 88.3 |
| 4 | 95.3 | 93.6 | 94.0 | 90.0 |
| 5 | 98.3 | 38.0 | 78.8 | 78.1 |
| 6 | 98.4 | 98.2 | 0 | 0 |
| 7 | 97.7 | 97.7 | 98.0 | 96.7 |

Jet flavour tagging

- Performance for flavour tagging are taken from note by Franco, Loukas and Michele

Jet Flavour Tagging for Future Colliders with Fast Simulation

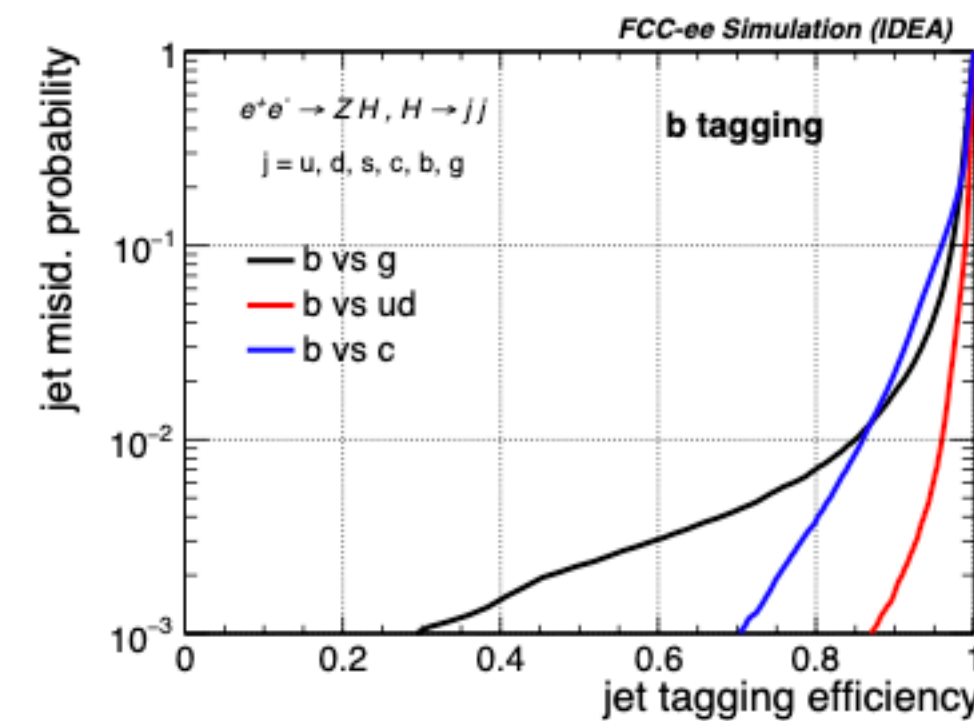
Franco Bedeschi^a, Loukas Gouskos^b and Michele Selvaggi^b

^aINFN Sezione di Pisa, Italy

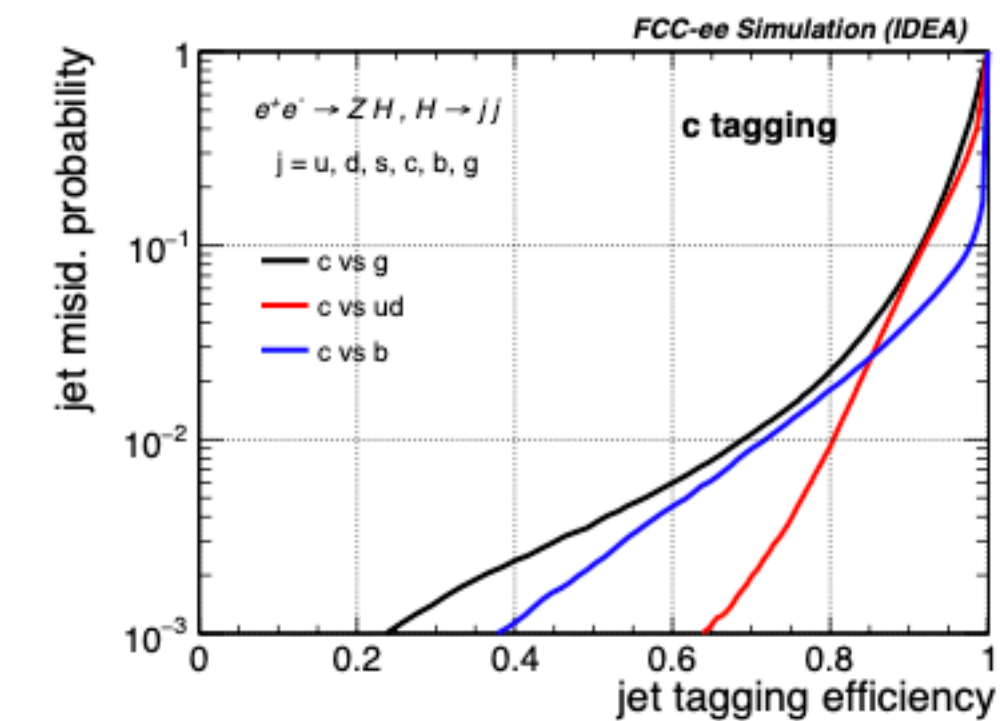
^bCERN, CH-1211 Geneva 23, Switzerland

E-mail: bed@fnal.gov, loukas.gouskos@cern.ch, michele.selvaggi@cern.ch

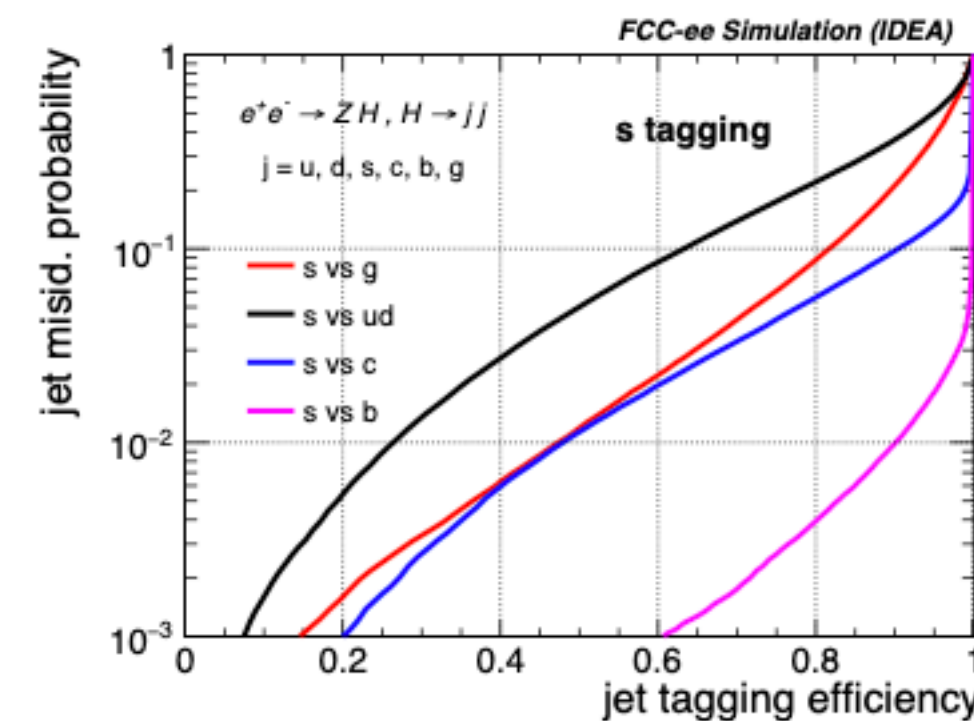
ABSTRACT: Jet flavour identification algorithms are of paramount importance to maximise the physics potential of future collider experiments. This work describes a novel set of tools allowing for a realistic simulation and reconstruction of particle level observables that are necessary ingredients to jet flavour identification. An algorithm for reconstructing the track parameters and covariance matrix of charged particles for an arbitrary tracking sub-detector geometries has been developed. Additional modules allowing for particle identification using time-of-flight and ionizing energy loss information have been implemented. A jet flavour identification algorithm based on a graph neural network architecture and exploiting all available particle level information has been developed. The impact of different detector design assumptions on the flavour tagging performance is assessed using the FCC-ee IDEA detector prototype.



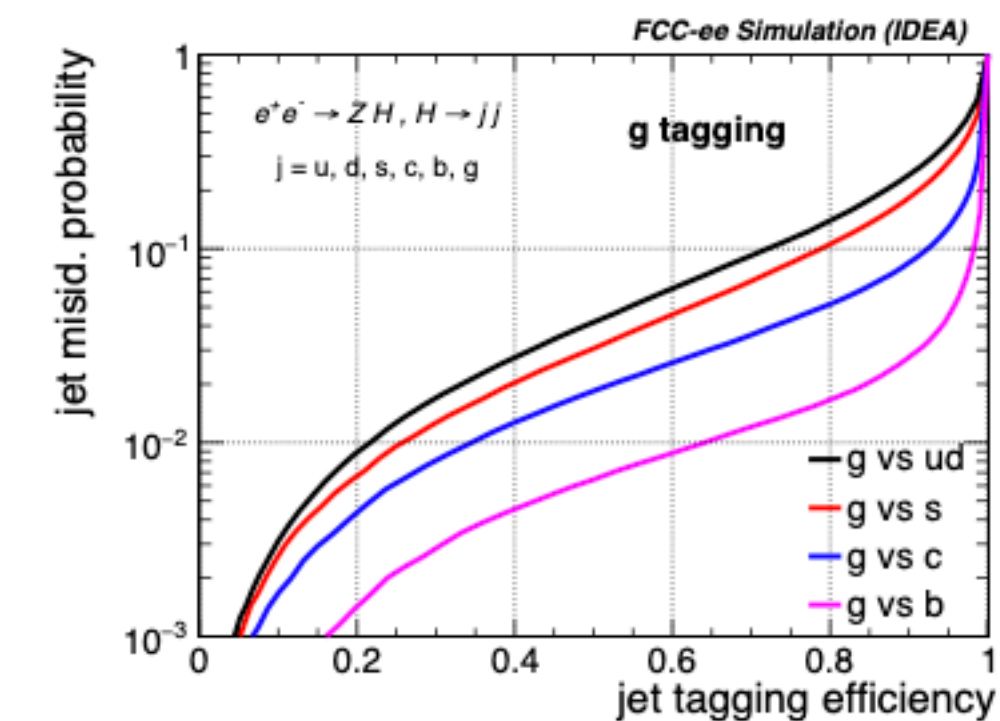
(a)



(b)



(c)



(d)

- For nominal analysis, use 80% efficiency working point for b, c, g tagging
- For alternative analyses, either vary efficiency for nominal fake rate or viceversa, or choose alternative WP (see next slide)

arXiv:2202.03285v1 [hep-ex] 7 Feb 2022

Jet flavour tagging (2)

- Flavour tagging scenarios considered:

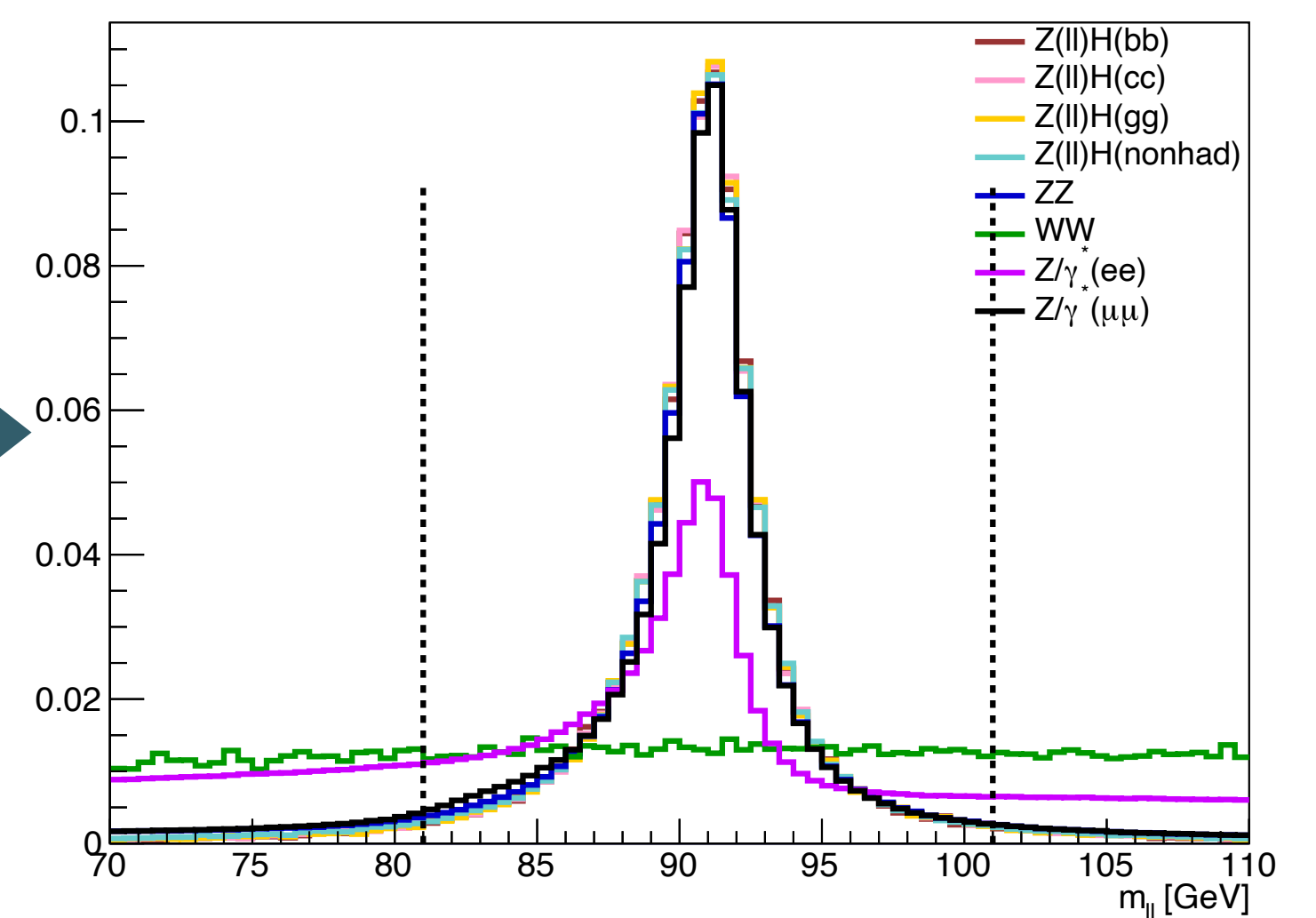
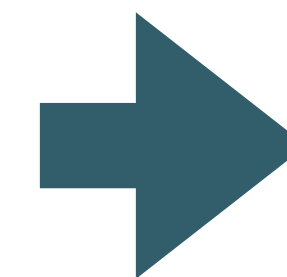
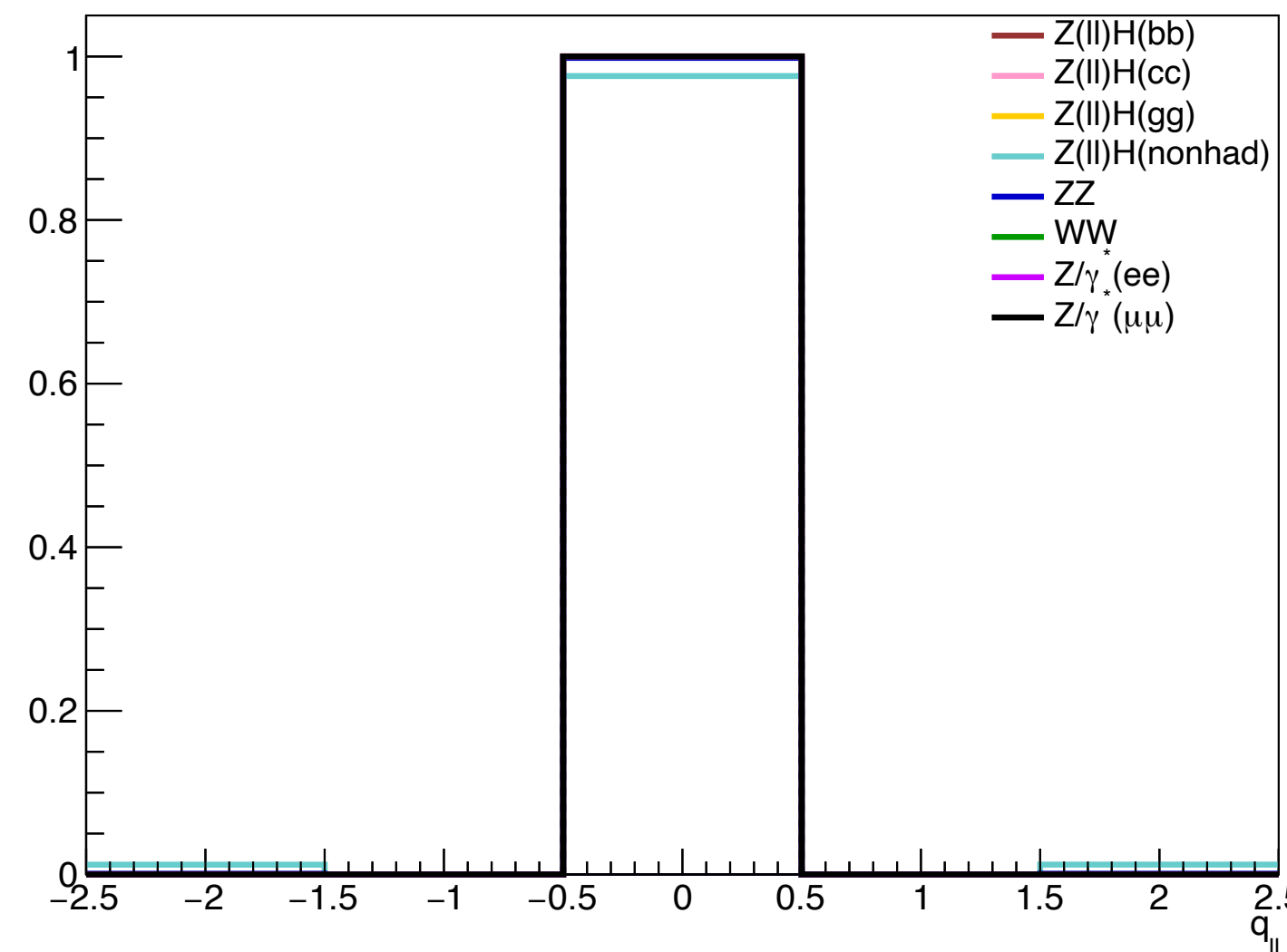
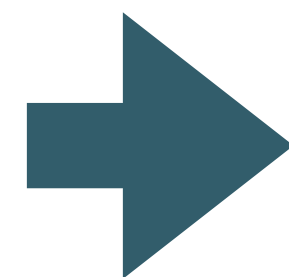
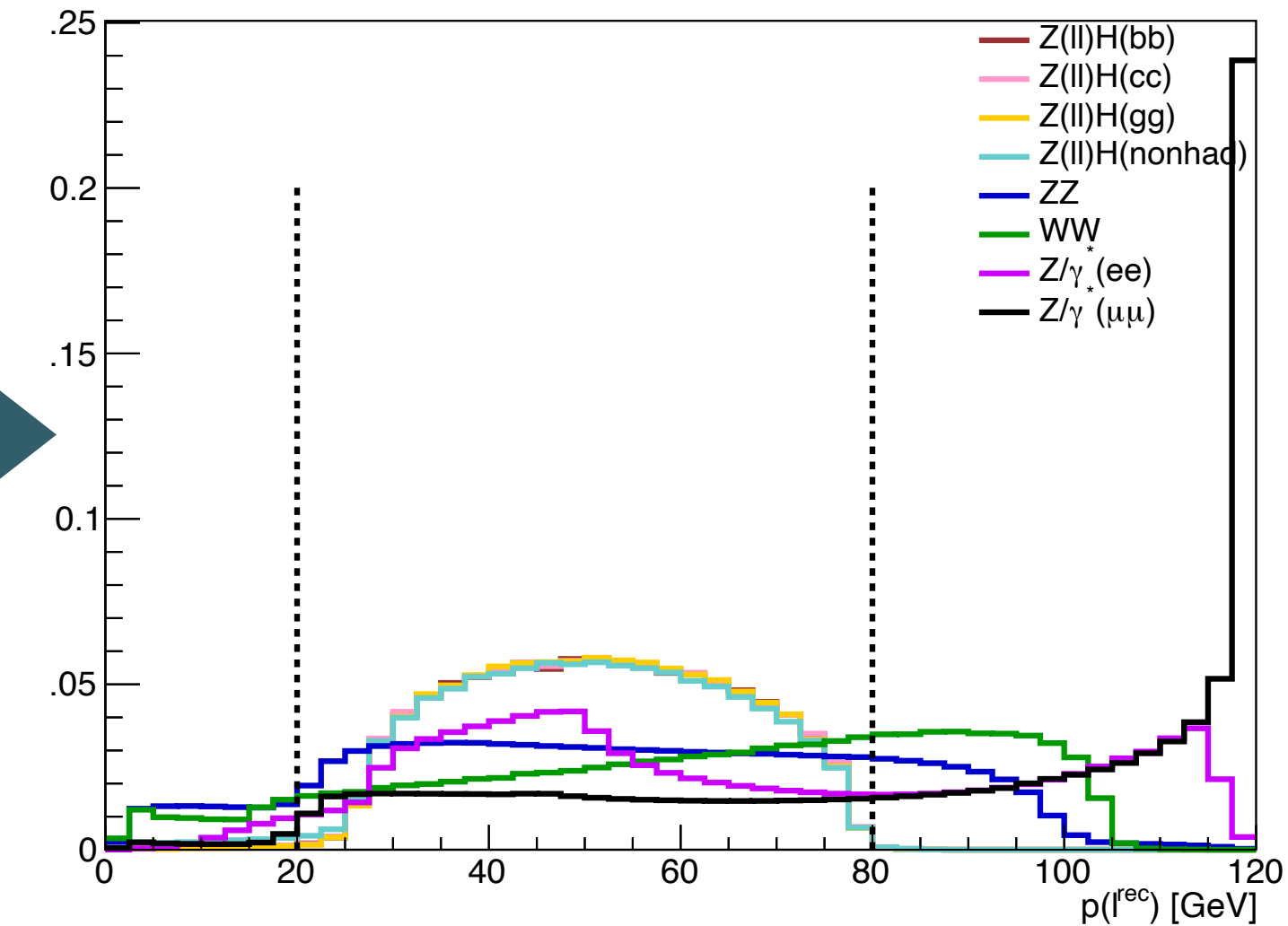
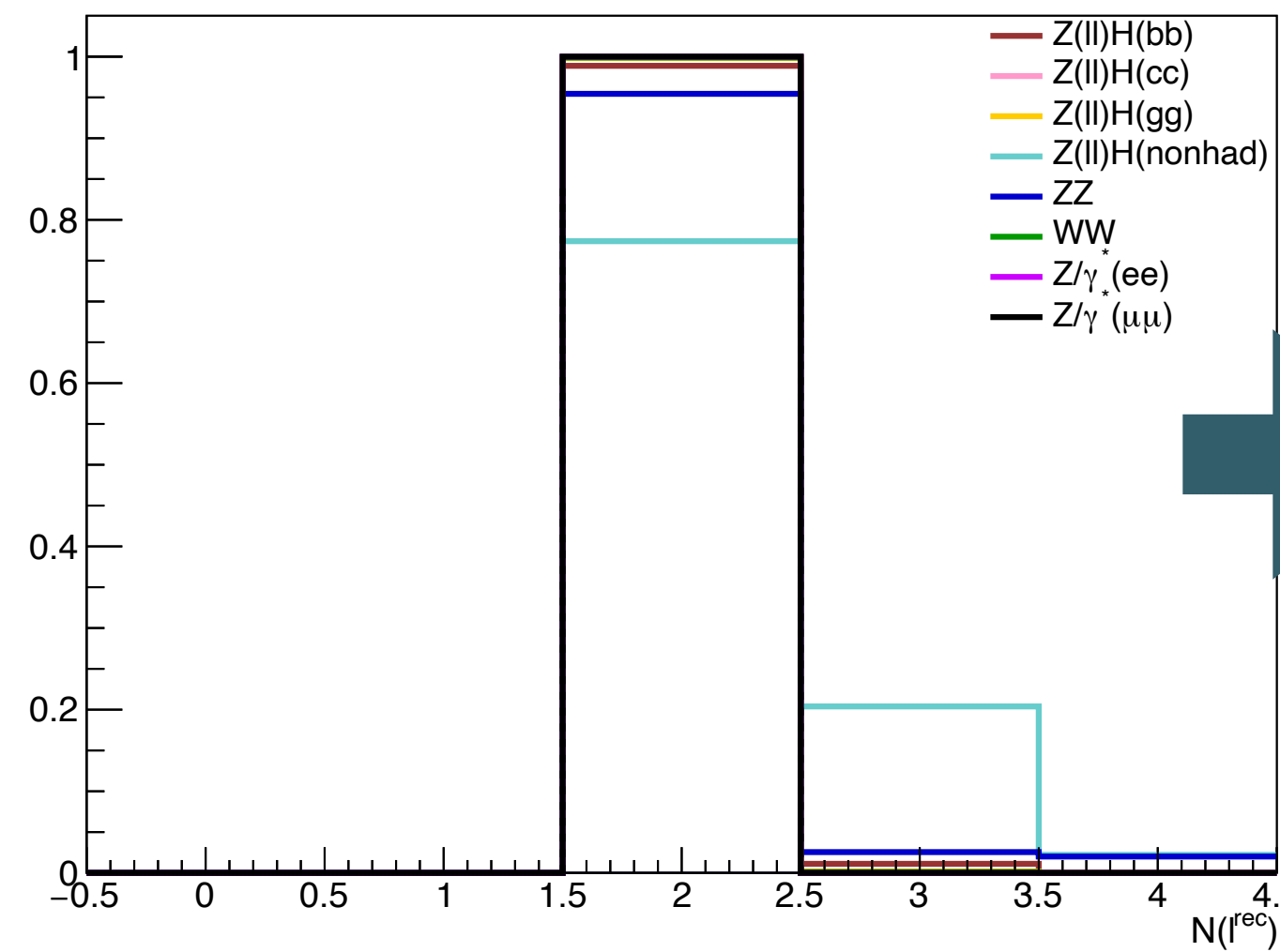
| Strategy | b-tag $\epsilon_b, \epsilon_c, \epsilon_l, \epsilon_g$ | c-tag $\epsilon_b, \epsilon_c, \epsilon_l, \epsilon_g$ | g-tag $\epsilon_b, \epsilon_c, \epsilon_l, \epsilon_g$ |
|---------------|---|---|---|
| Nominal | 80 / 0.4 / 0.05 / 0.7 | 2.0 / 80 / 0.9 / 2.5 | 2.0 / 5.0 / 15 / 80 |
| Fake rates x2 | 80 / 0.8 / 0.1 / 1.4 | 4.0 / 80 / 1.8 / 5.0 | 4.0 / 10 / 30 / 80 |
| Fake rates x5 | 80 / 2.0 / 0.25 / 3.5 | 10 / 80 / 4.5 / 12.5 | 10 / 25 / 75 / 80 |
| Eff -10% | 70 / 0.4 / 0.05 / 0.7 | 2.0 / 70 / 0.9 / 2.5 | 2.0 / 5.0 / 15 / 70 |
| Eff -20% | 60 / 0.4 / 0.05 / 0.7 | 2.0 / 60 / 0.9 / 2.5 | 2.0 / 5.0 / 15 / 60 |
| WPc 90% | 80 / 0.4 / 0.05 / 0.7 | 4.0 / 90 / 7.0 / 7.0 | 2.0 / 5.0 / 15 / 80 |
| WPc 70% | 80 / 0.4 / 0.05 / 0.7 | 0.9 / 70 / 0.2 / 1.0 | 2.0 / 5.0 / 15 / 80 |

- For alternative WPs, we investigated the case of changing the c-tagging one since BR(cc) is the one which is projected to have the less precise measurement (see later)

Event selection

- **Z->ll selection:**

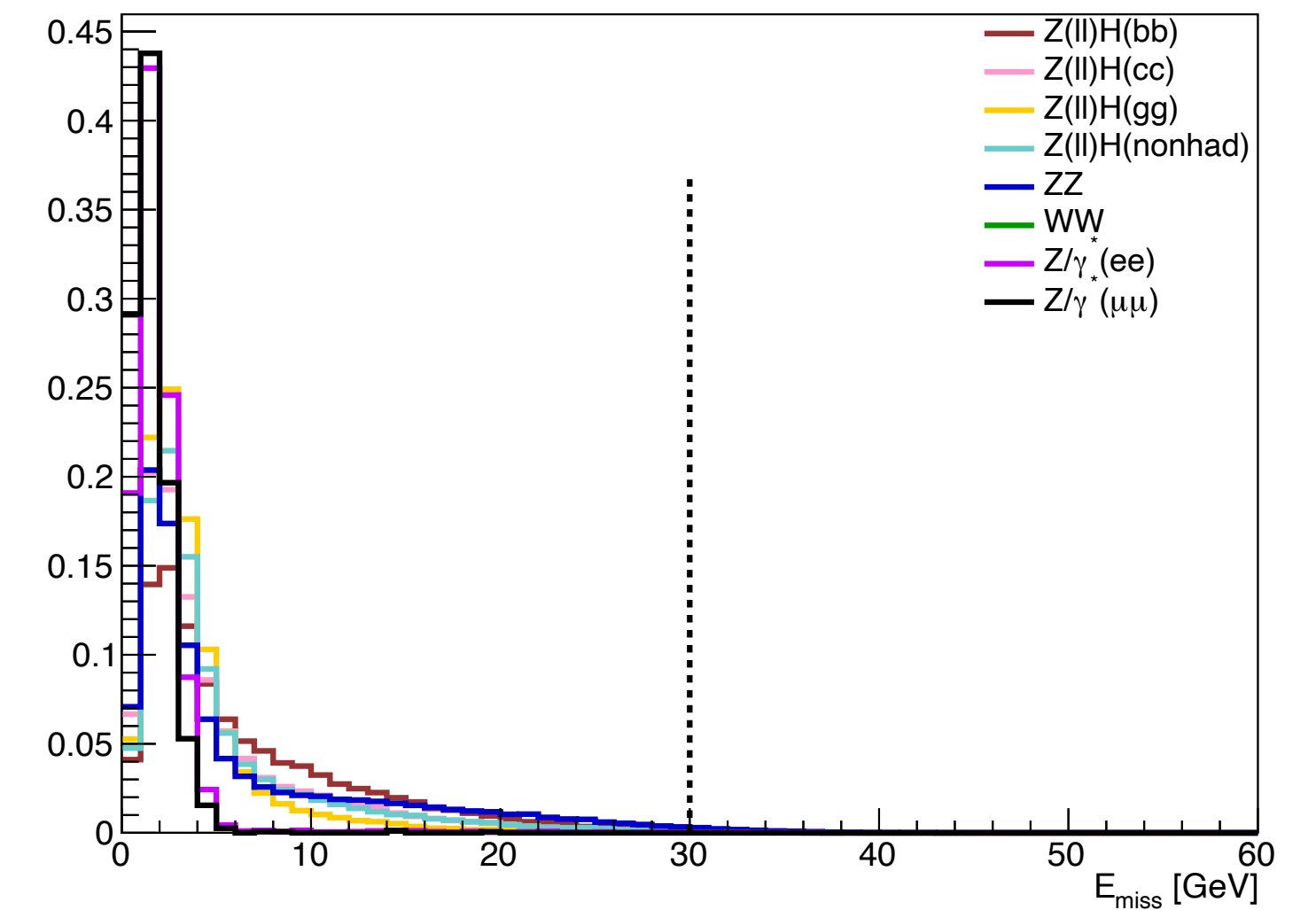
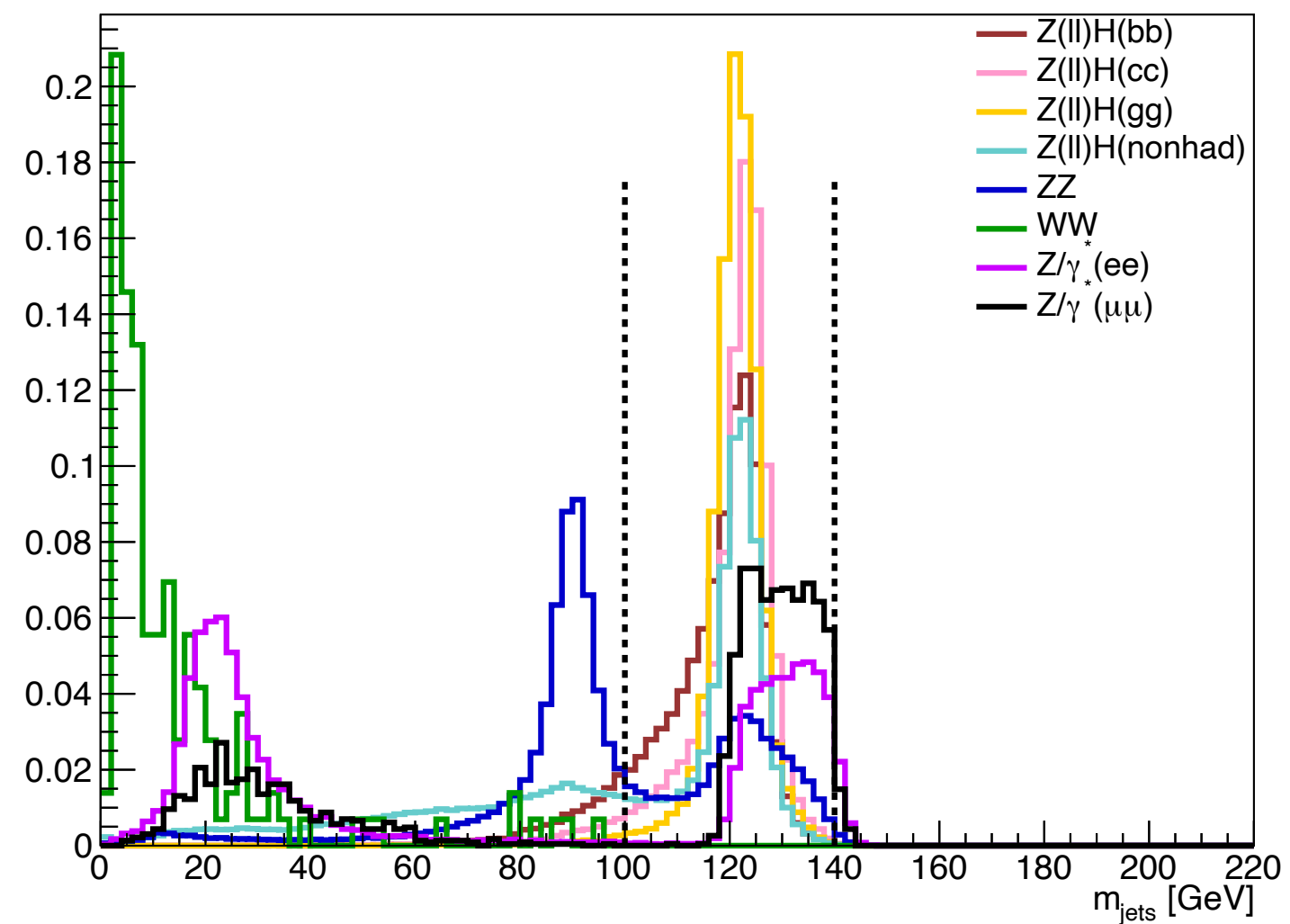
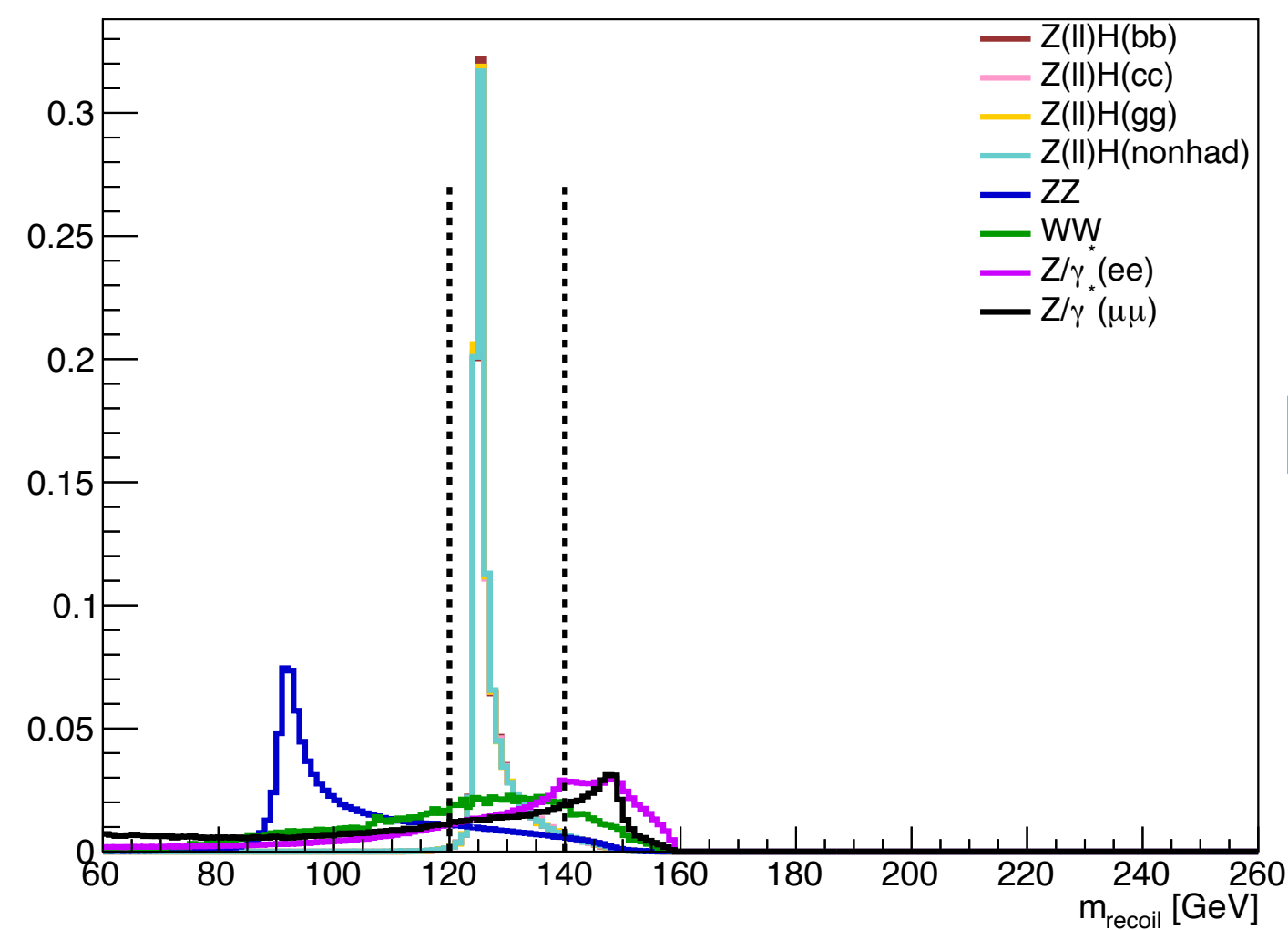
- Exactly 2 isolated electrons or muons
- Lepton momenta between 25 and 80 GeV
- $Q(ll)=0$
- Dilepton invariant mass in 81-101 GeV
- $|\cos(\text{Polar angle of dilepton pair})| < 0.8$



Event selection (2)

- **Recoil and jet selection:**

- Recoil mass in 120-140 GeV
- Jet momentum in 10-100 GeV
- Hadronic mass in 100-140 GeV
- Missing energy < 30 GeV
- Tighter cuts (15, 20 GeV) in gluon-enriched and c-enriched categories



Cutflow, expected yields, efficiency

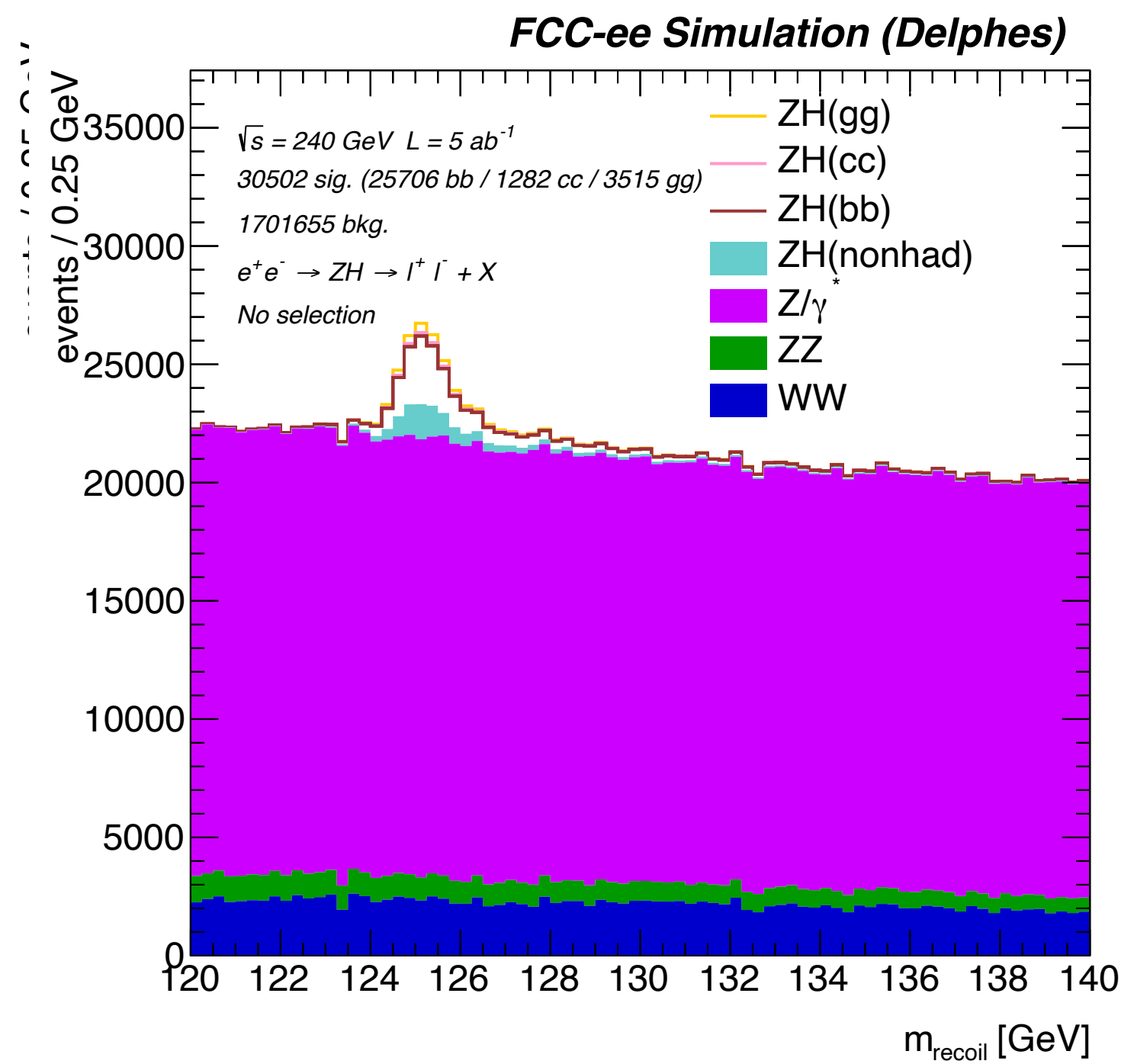
| Cut | ZHbb | | ZHcc | | ZHgg | | ZHnonhad | | ZZ | WW | Z/y*(ee) | Z/y*(mm) | Z/y*(qq) |
|-------------------------|-------|-----|-------|-----|-------|-----|----------|---------|----------|----------|----------|-----------|----------|
| | Yield | Sig | Yield | Sig | Yield | Sig | Yield | Yield | Yield | Yield | Yield | Yield | Yield |
| No cuts | 40750 | 2 | 2023 | 0 | 5730 | 0 | 21470 | 6794950 | 82192500 | 41525000 | 26440000 | 263269500 | |
| 2e or 2mu | 27045 | 4 | 1357 | 0 | 3724 | 1 | 12375 | 871502 | 2130832 | 29524745 | 18699307 | 45888 | |
| No extra lep | 26739 | 4 | 1356 | 0 | 3718 | 1 | 9573 | 831747 | 2128490 | 29524745 | 18699307 | 44940 | |
| p(lep) 25-80 GeV | 26124 | 8 | 1328 | 0 | 3646 | 1 | 9001 | 426207 | 700864 | 7172263 | 2633299 | 421 | |
| q(ll)=0 | 26122 | 8 | 1328 | 0 | 3646 | 1 | 8785 | 425439 | 700683 | 7172263 | 2633299 | 421 | |
| m(ll) 81-101 GeV | 24463 | 11 | 1244 | 1 | 3417 | 1 | 8049 | 295892 | 170319 | 2732241 | 1989103 | 0 | |
| cos(theta_ll) <0.8 | 19950 | 20 | 1017 | 1 | 2787 | 3 | 6593 | 187765 | 128845 | 440918 | 242606 | 0 | |
| m(recoil) 120-140 GeV | 19186 | 38 | 979 | 2 | 2677 | 5 | 6318 | 31090 | 53376 | 106466 | 31776 | 0 | |
| 100<m(jets)<140 GeV | 17103 | 85 | 932 | 5 | 2637 | 13 | 3804 | 9397 | 0 | 6236 | 830 | 0 | |
| Emiss < 30 GeV | 17063 | 85 | 931 | 5 | 2637 | 13 | 3796 | 9283 | 0 | 6226 | 830 | 0 | |
| per-category Emiss cuts | 16992 | 85 | 909 | 5 | 2593 | 13 | 3615 | 8377 | 0 | 6180 | 830 | 0 | |

| Efficiency (%) | ZHbb | ZHcc | ZHgg | ZHnonhad | WW | ZZ | Zqq | Zee | Zmumu |
|----------------|-------|-------|-------|----------|------|------|------|------|-------|
| | 41.70 | 44.95 | 45.25 | 16.84 | 0.00 | 0.12 | 0.00 | 0.01 | 0.00 |

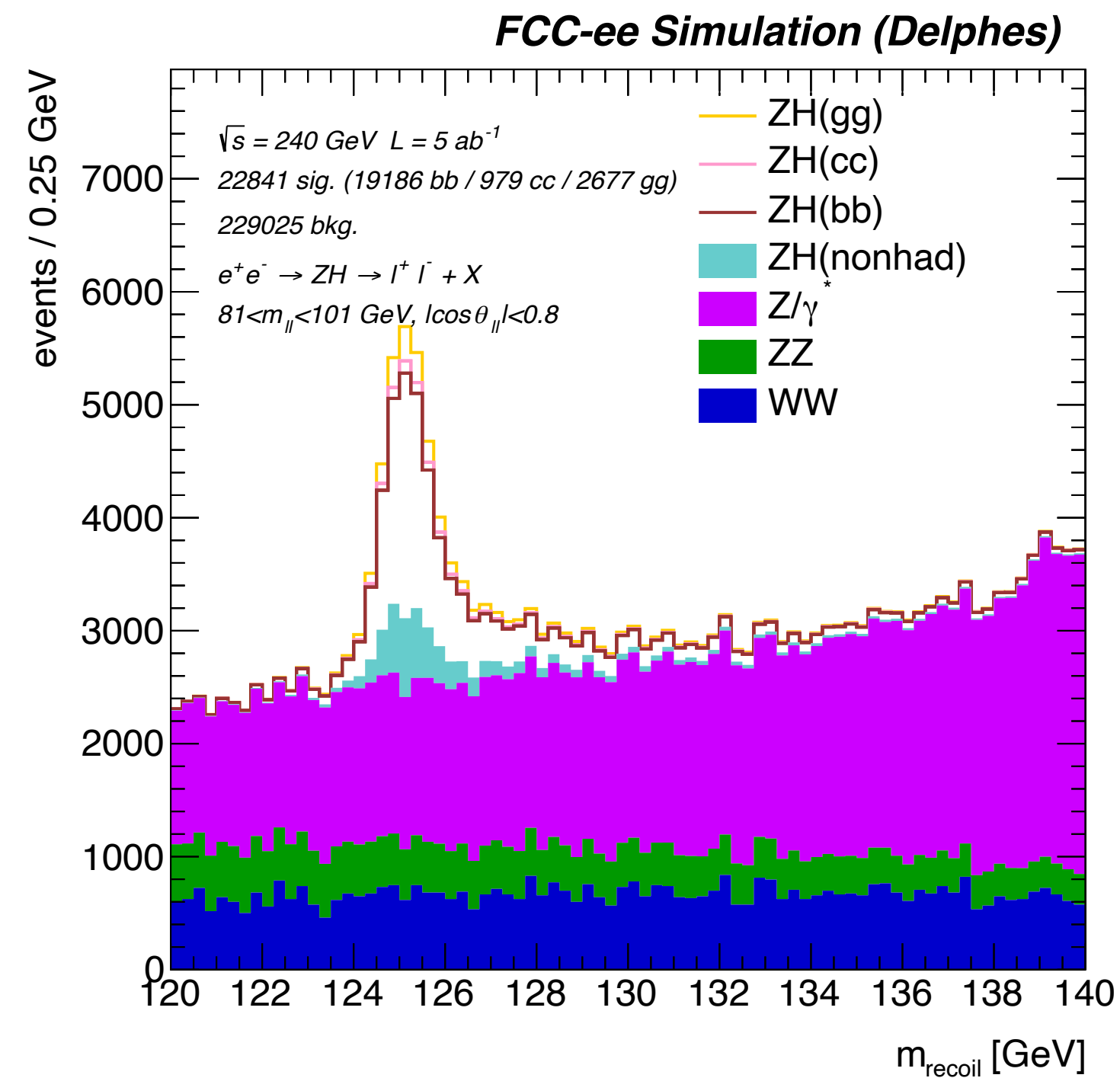
- 2e or 2mu request leads to 65% efficiency despite very high lepton reconstruction efficiency due to isolation requirement inefficiency

Evolution of the m_{recoil} distribution after the selection steps

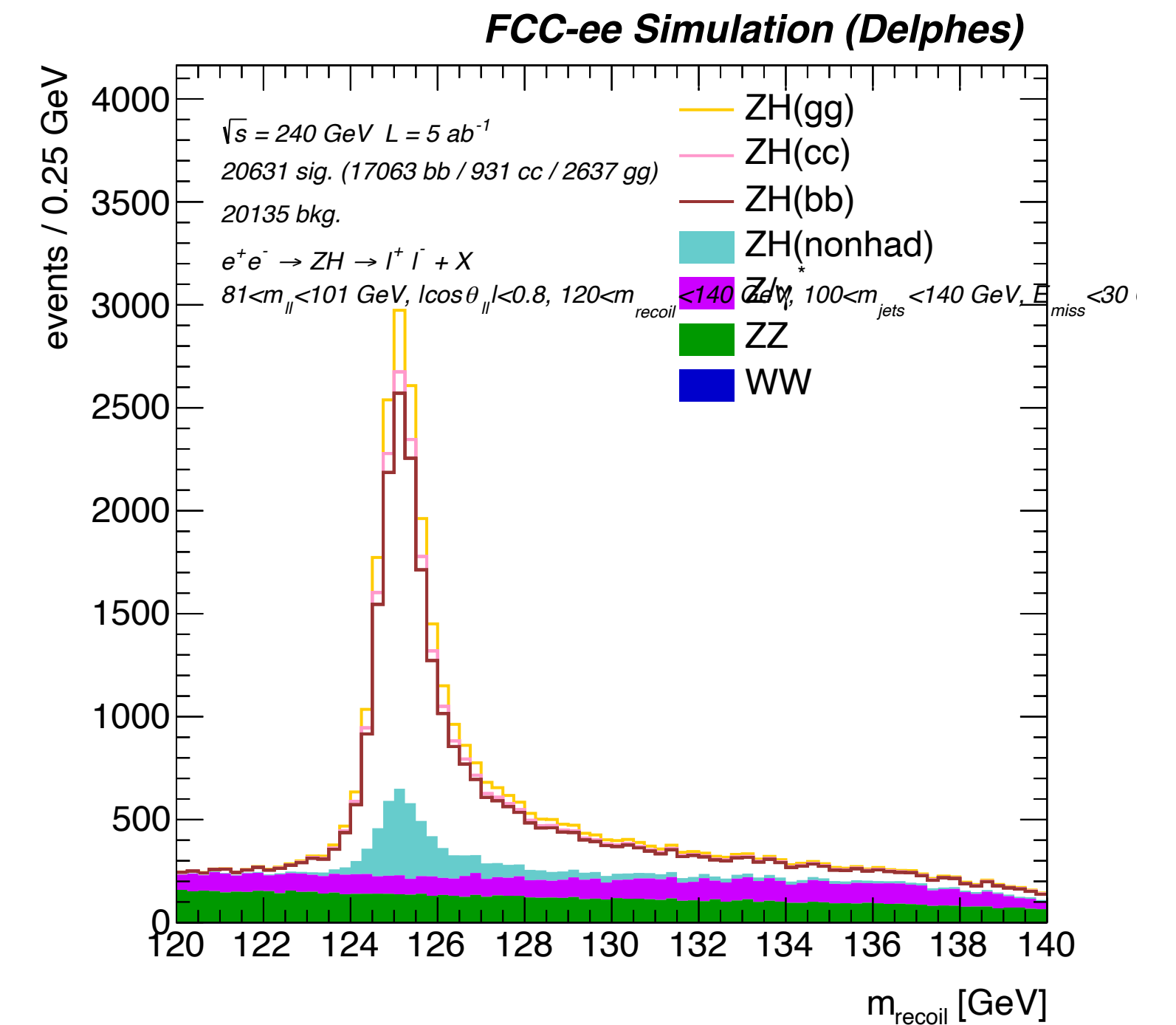
After reconstruction



After Z selection

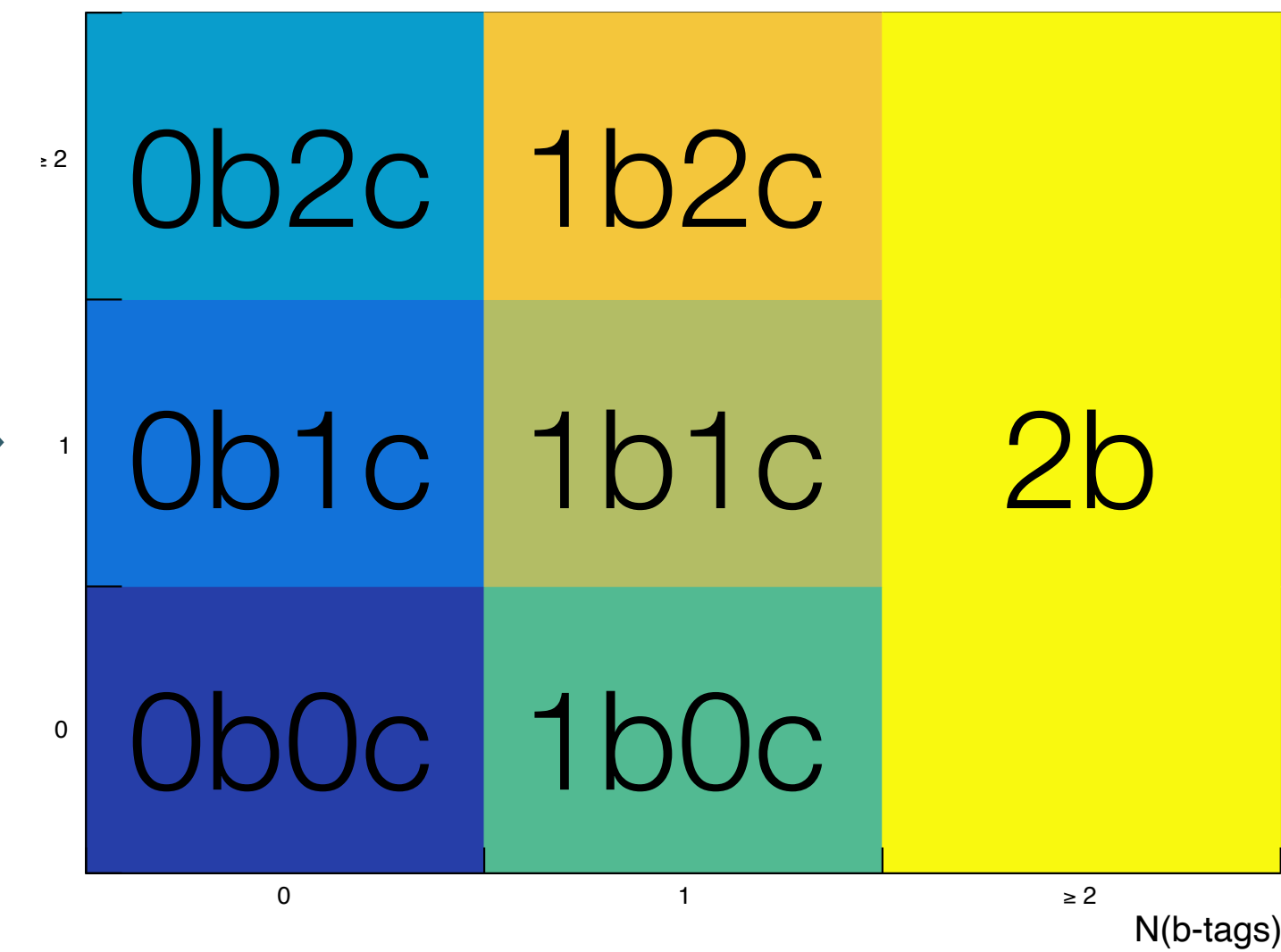
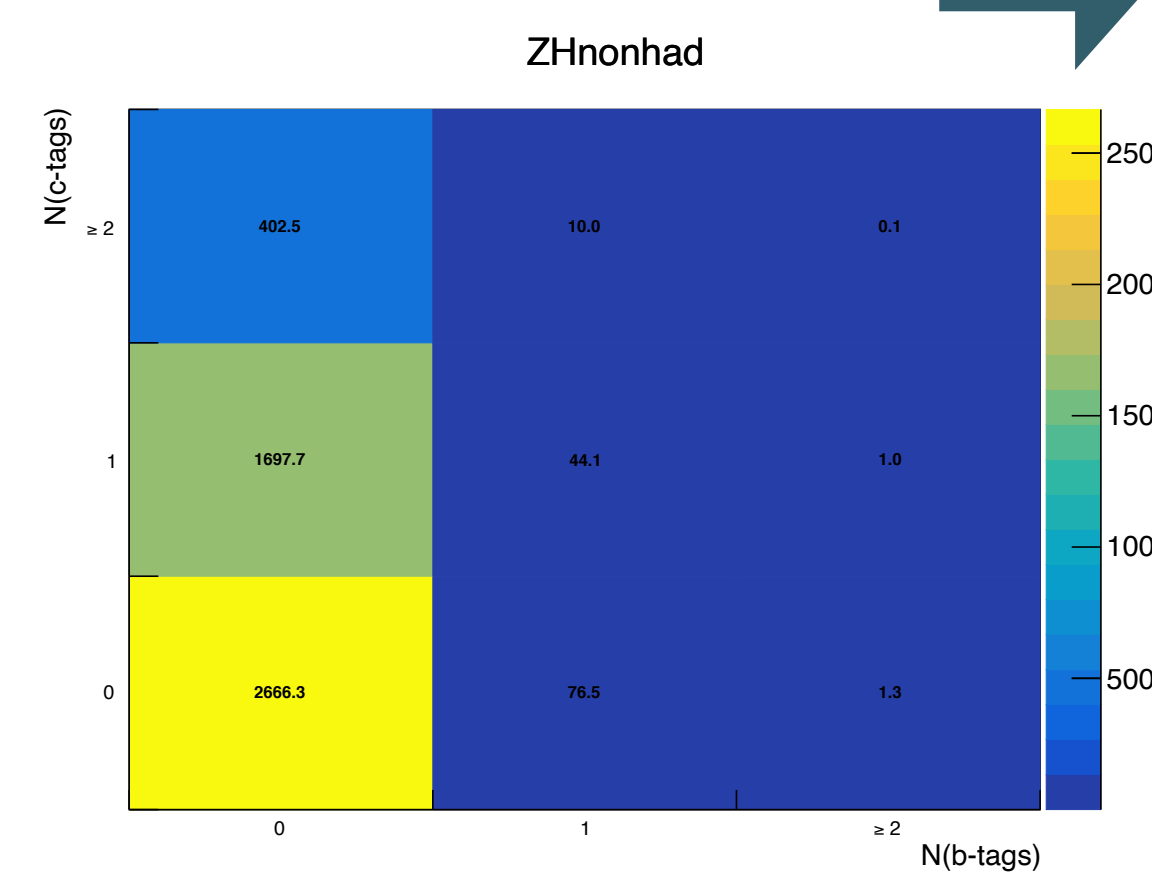
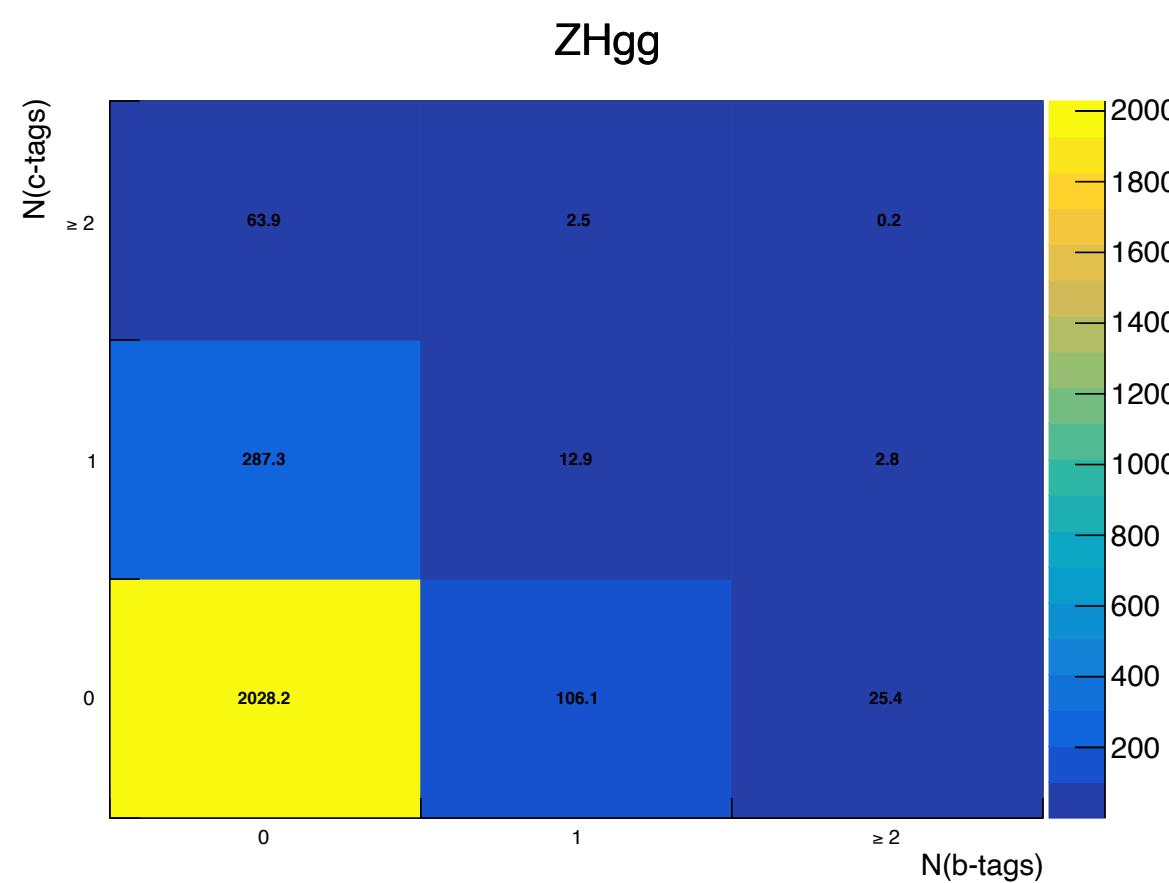
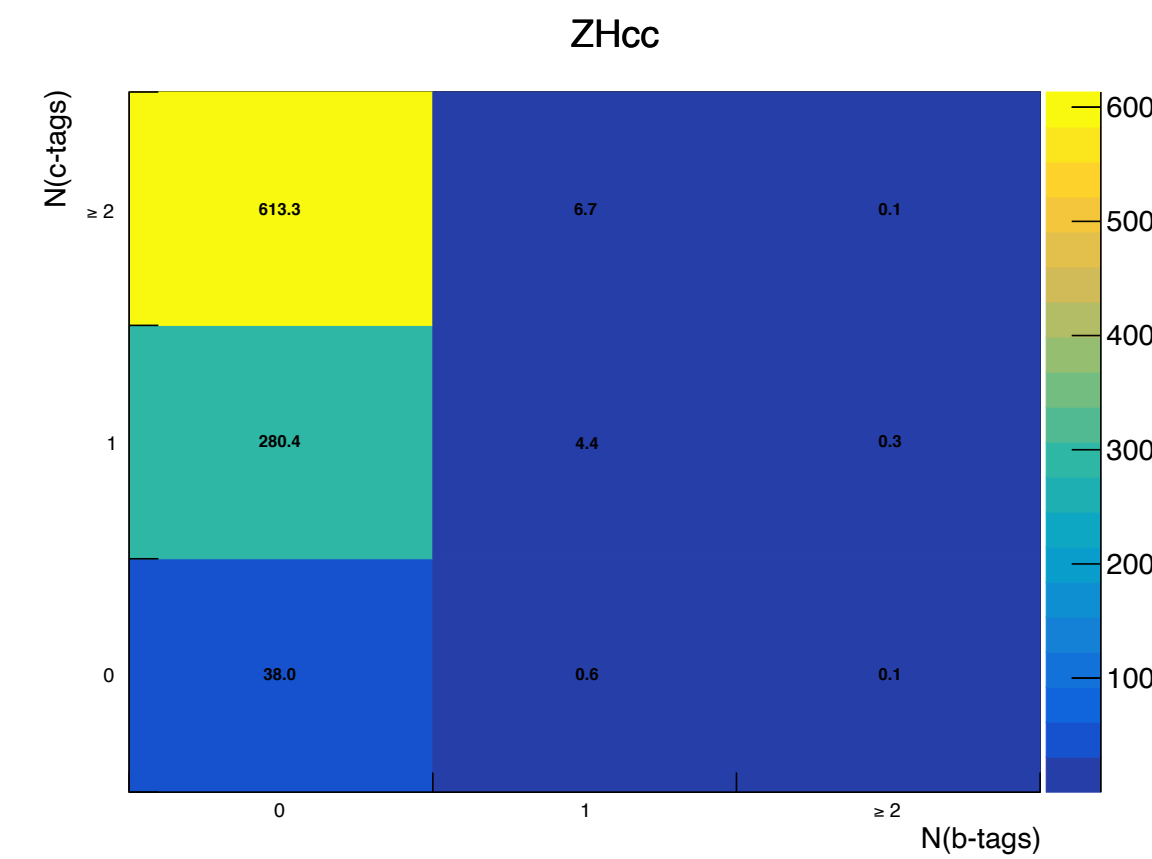
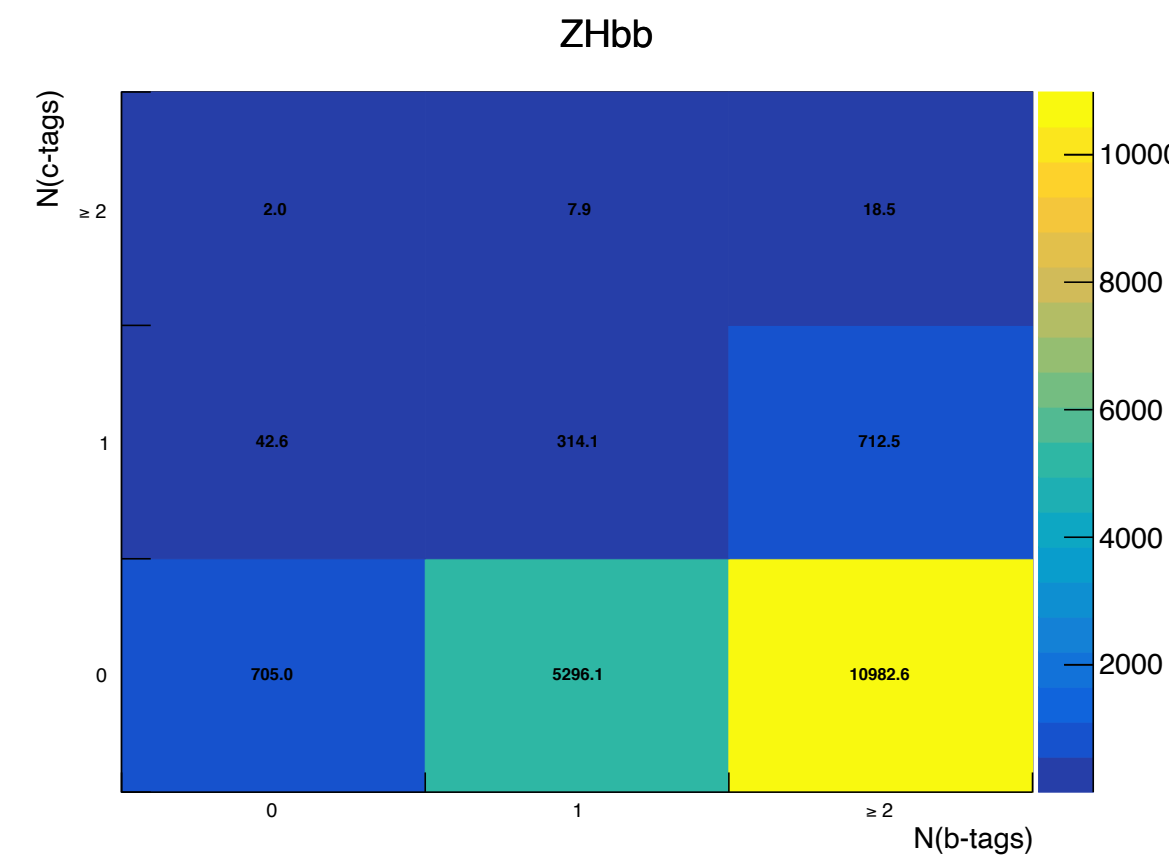
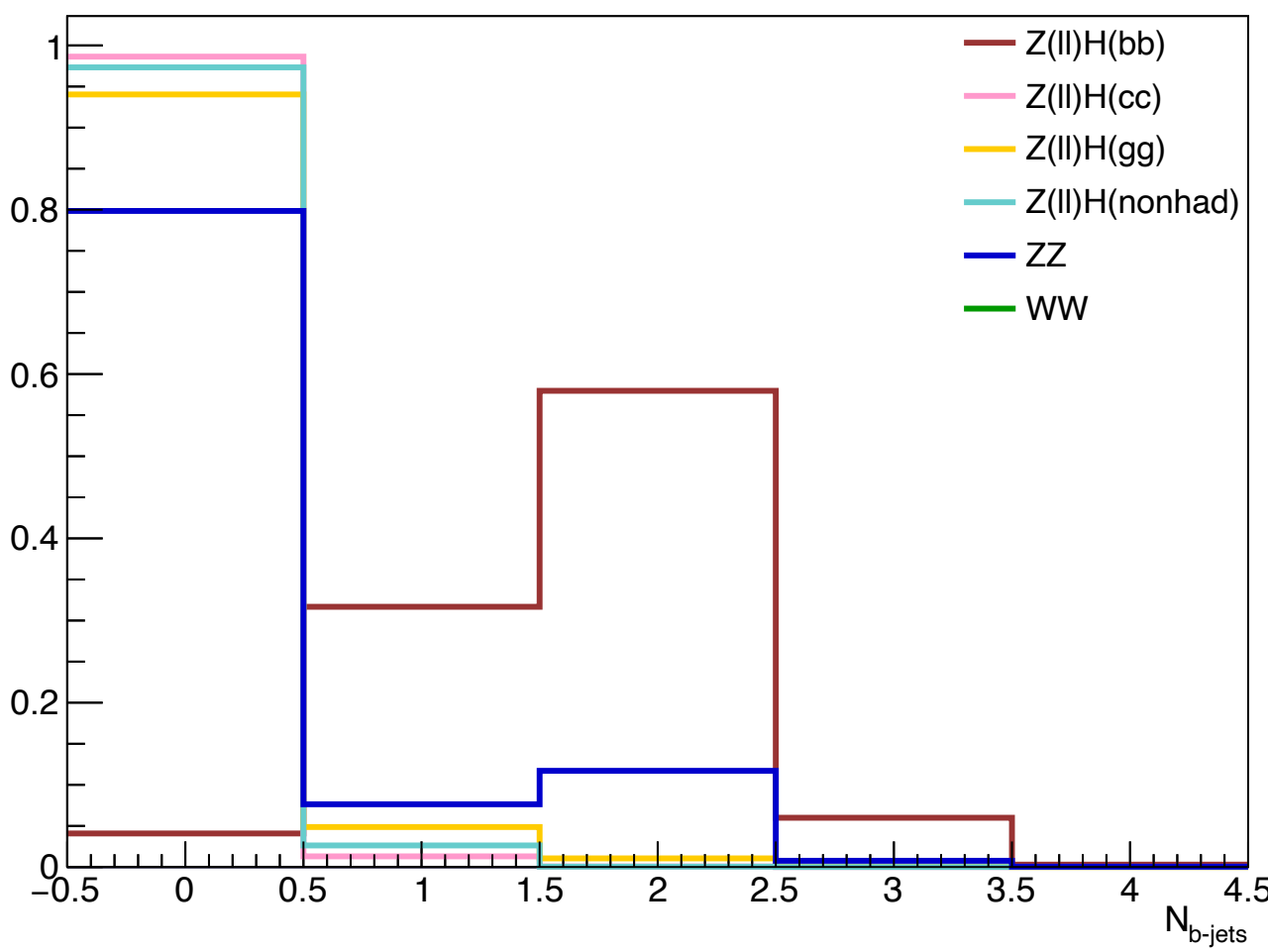


After recoil, jets, E_{miss} selection



Event categorisation

- Events are classified in mutually orthogonal categories based on the number of b-, c- and g-tags



Event categorisation

- Events are classified in mutually orthogonal categories based on the number of b-, c- and g-tags (2c category also requires < 2b tags)

EXPECTED YIELDS (significances in parentheses)

| | ZHbb | ZHcc | ZHgg | ZHnonhad | bkg |
|--------|-------------|----------|-----------|-----------|-------|
| 2b | 10777 (100) | 0 (0) | 1 (0) | 46 (0) | 814 |
| 2c | 2 (0) | 577 (15) | 2 (0) | 66 (2) | 799 |
| 1b1c | 219 (13) | 2 (0) | 2 (0) | 10 (1) | 34 |
| 1b0c | 5378 (69) | 0 (0) | 39 (0) | 40 (1) | 680 |
| 0b1c | 33 (1) | 292 (6) | 131 (3) | 1193 (24) | 926 |
| 0b0c2g | 4 (0) | 1 (0) | 1533 (34) | 50 (1) | 392 |
| 0b0c1g | 60 (1) | 5 (0) | 781 (11) | 574 (8) | 3614 |
| 0b0c0g | 518 (5) | 31 (0) | 103 (1) | 1636 (16) | 8129 |
| Total | 16992 (122) | 909 (16) | 2593 (36) | 3615 (30) | 15386 |

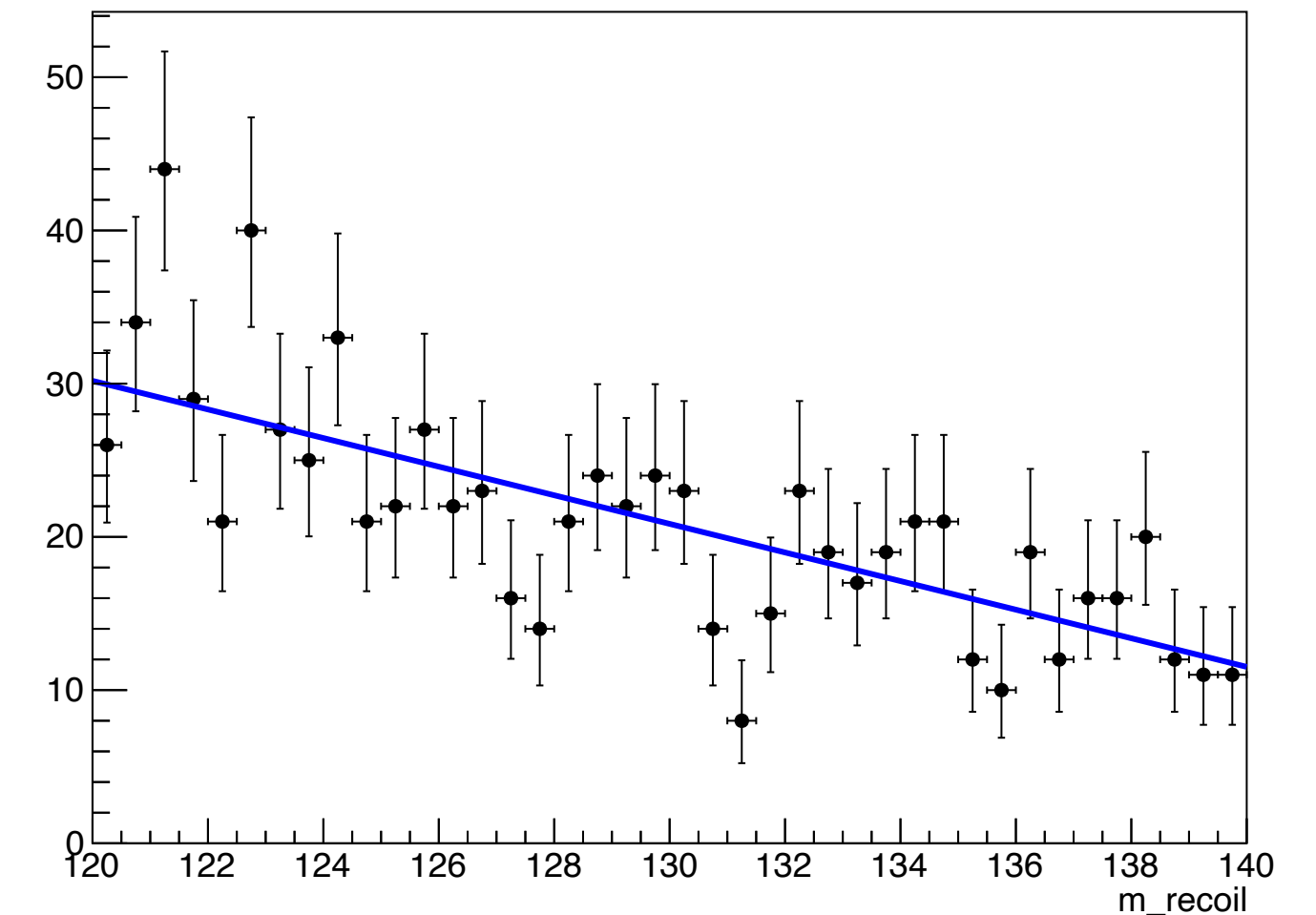
Fit - likelihood model

- **Simultaneous S+B fit to the recoil mass** of the event categories
- **Background** model: simple functions (**polynomials**, exponentials) with **floating parameters** in each category
- **Signal** model: double-sided **Crystal Ball** with same parameters in each category
 - **peak position = mH + constant** (checked with MC samples w/ different mH)
 - Tail parameters and peak-mH are fixed, **m_H and resolution are floating**
- Signal yield in each category = function of the efficiencies for the various Z(ll)H(->XX) processes in each category (fixed, from simulation) and of $\sigma_H \cdot BR(H \rightarrow XX)$:

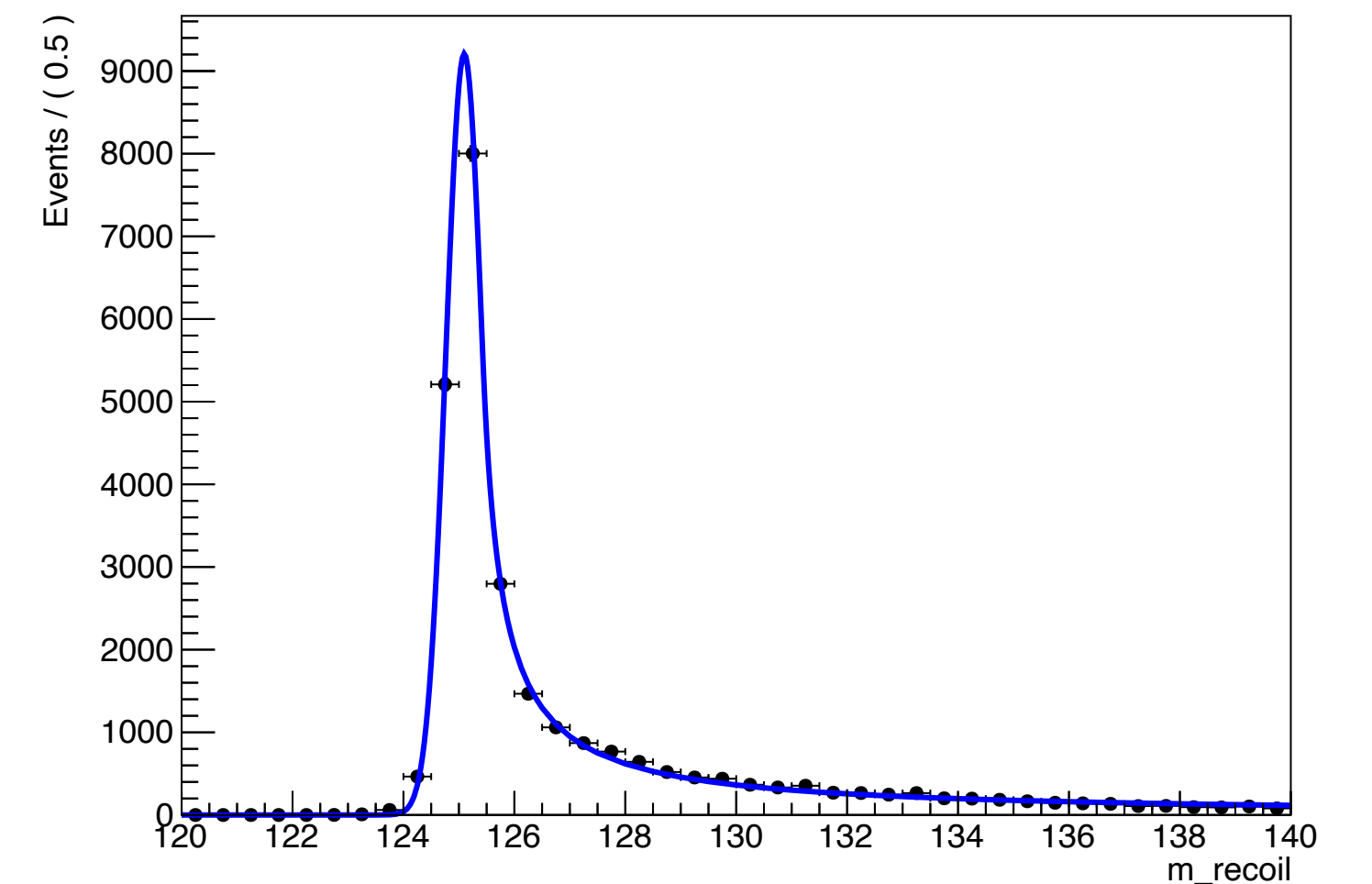
$$N_i = L \times \sigma(ee \rightarrow ZH) \times BR(Z \rightarrow ll) \times \left(BR(H \rightarrow b\bar{b})\epsilon_i^{b\bar{b}} + BR(H \rightarrow c\bar{c})\epsilon_i^{c\bar{c}} + BR(H \rightarrow gg)\epsilon_i^{gg} + BR(H \rightarrow nonhad)\epsilon_i^{nh} \right).$$

- In fit, $\sigma \cdot BR(H \rightarrow XX) = (\sigma \cdot BR(H \rightarrow XX))_{SM} \cdot K_{XX} \Rightarrow$ **parameters of interest = {K_{XX}}**
- Fit the output of the simulation (not an Asimov sample generated from the nominal models) \Rightarrow statistical deviations from expected values $k_{XX}=1$ are expected
- The fit is **binned**, in the m_{recoil} range **120-140 GeV**
- **Non-hadronic BR is fixed to SM prediction (assume to measure it precisely with other channels) - though some constraining power from 0b0c0g category**

ee -> ZZ 2b

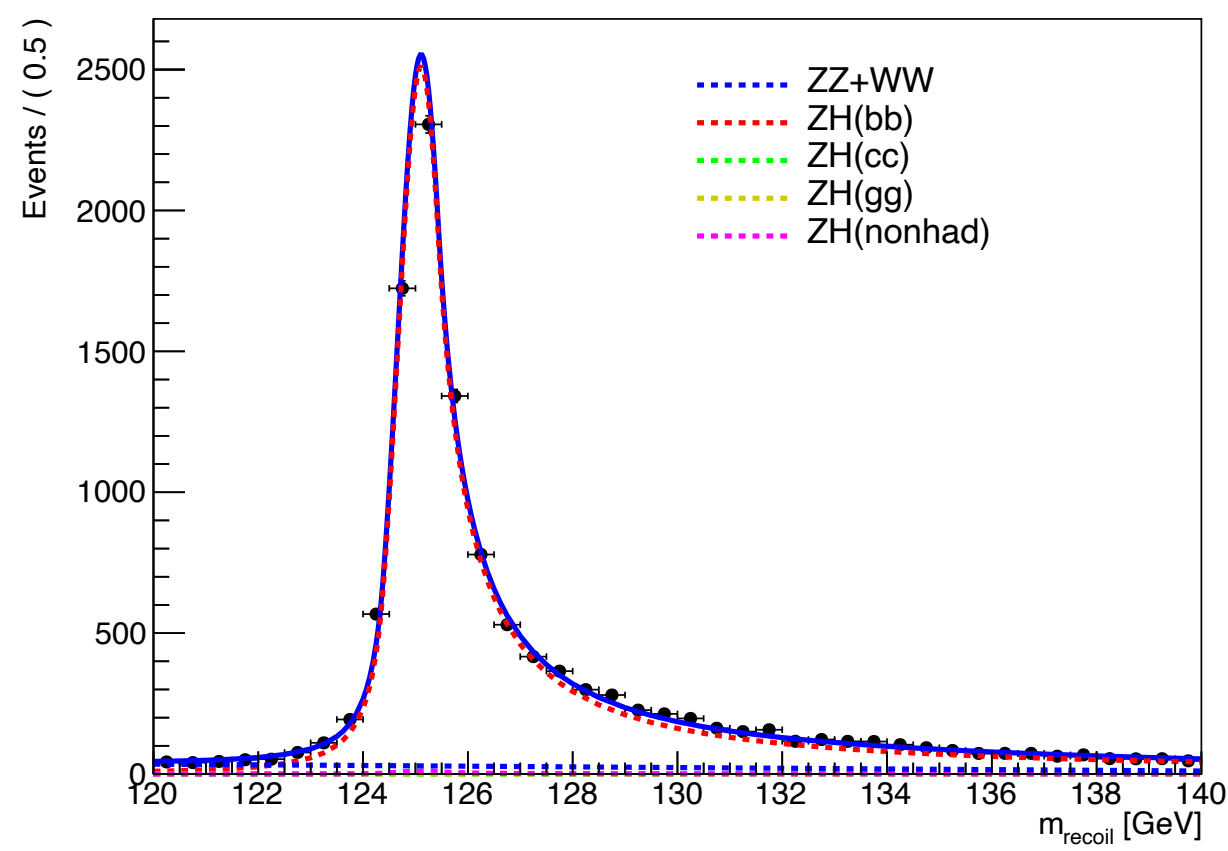


ee -> ZHbb 2b

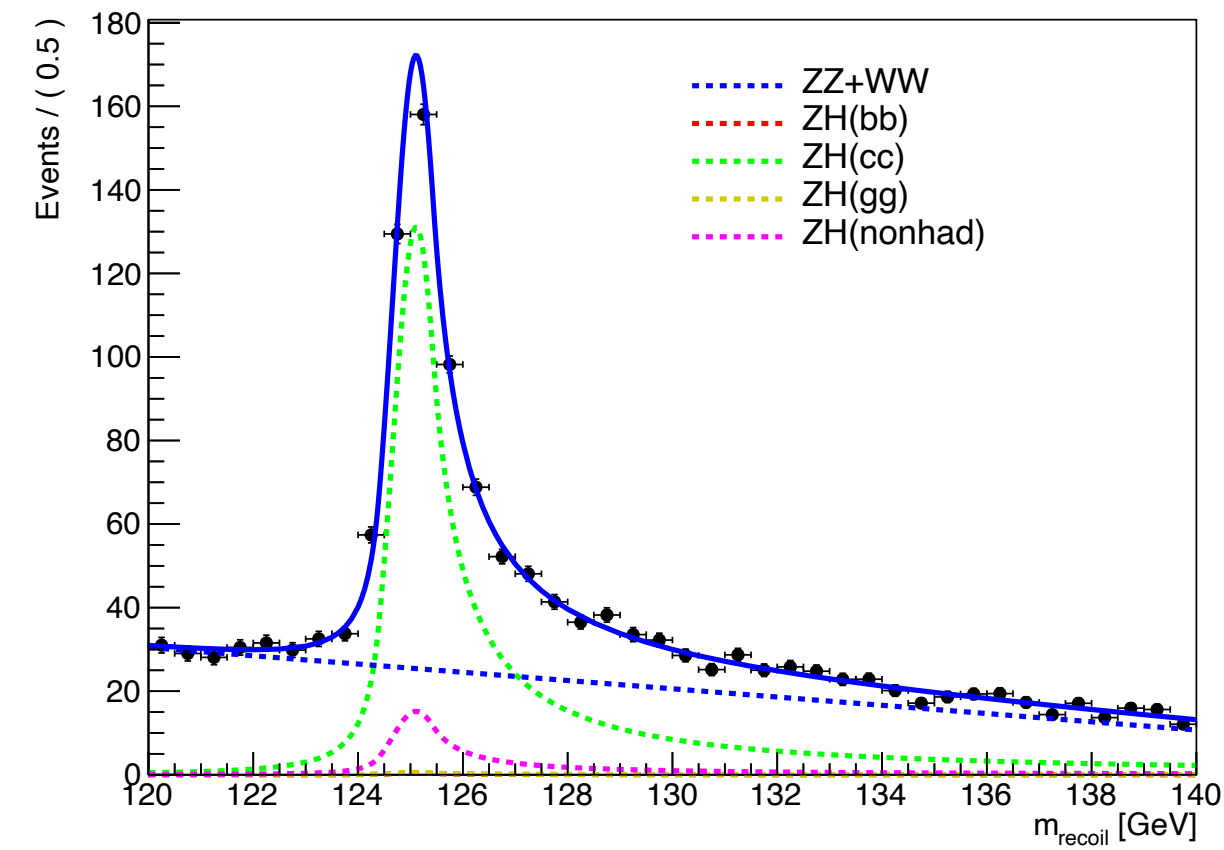


Results - nominal selection

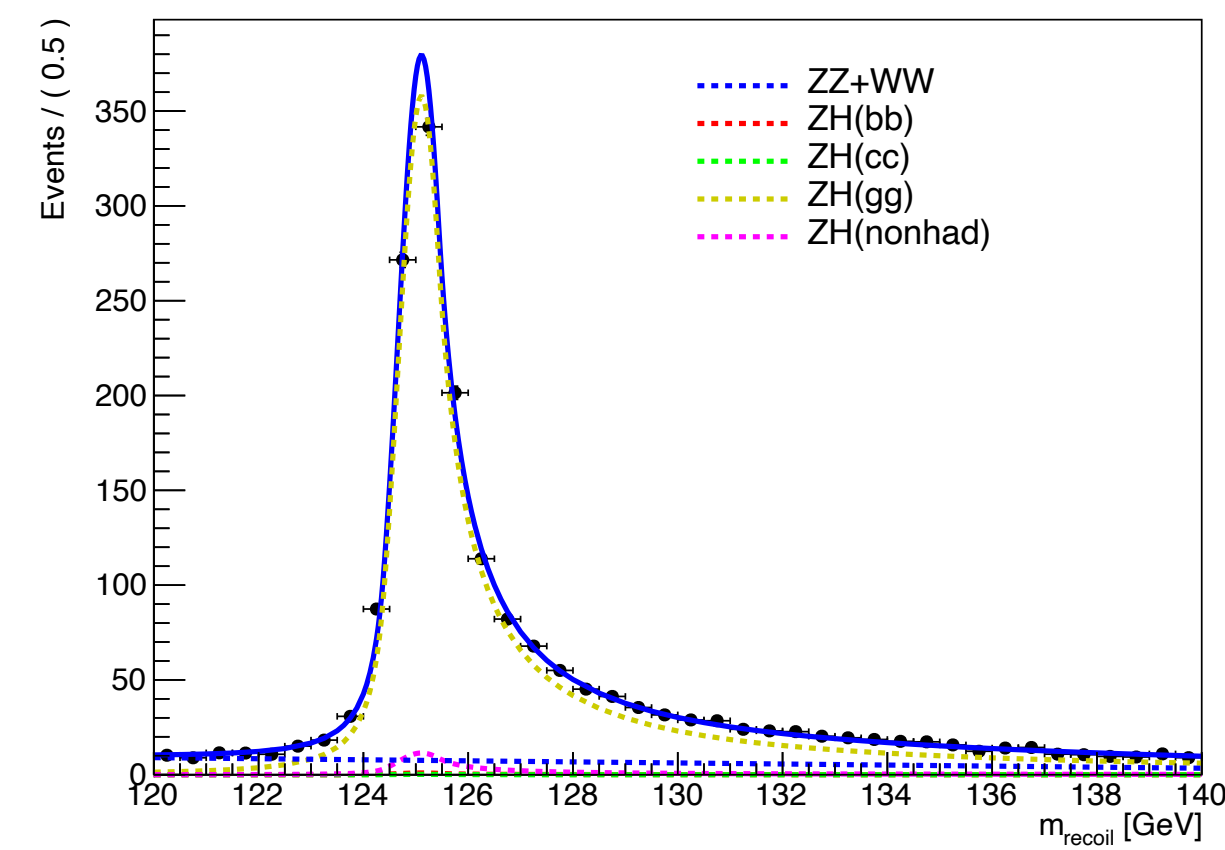
ee -> ZH, WW, ZZ 2b



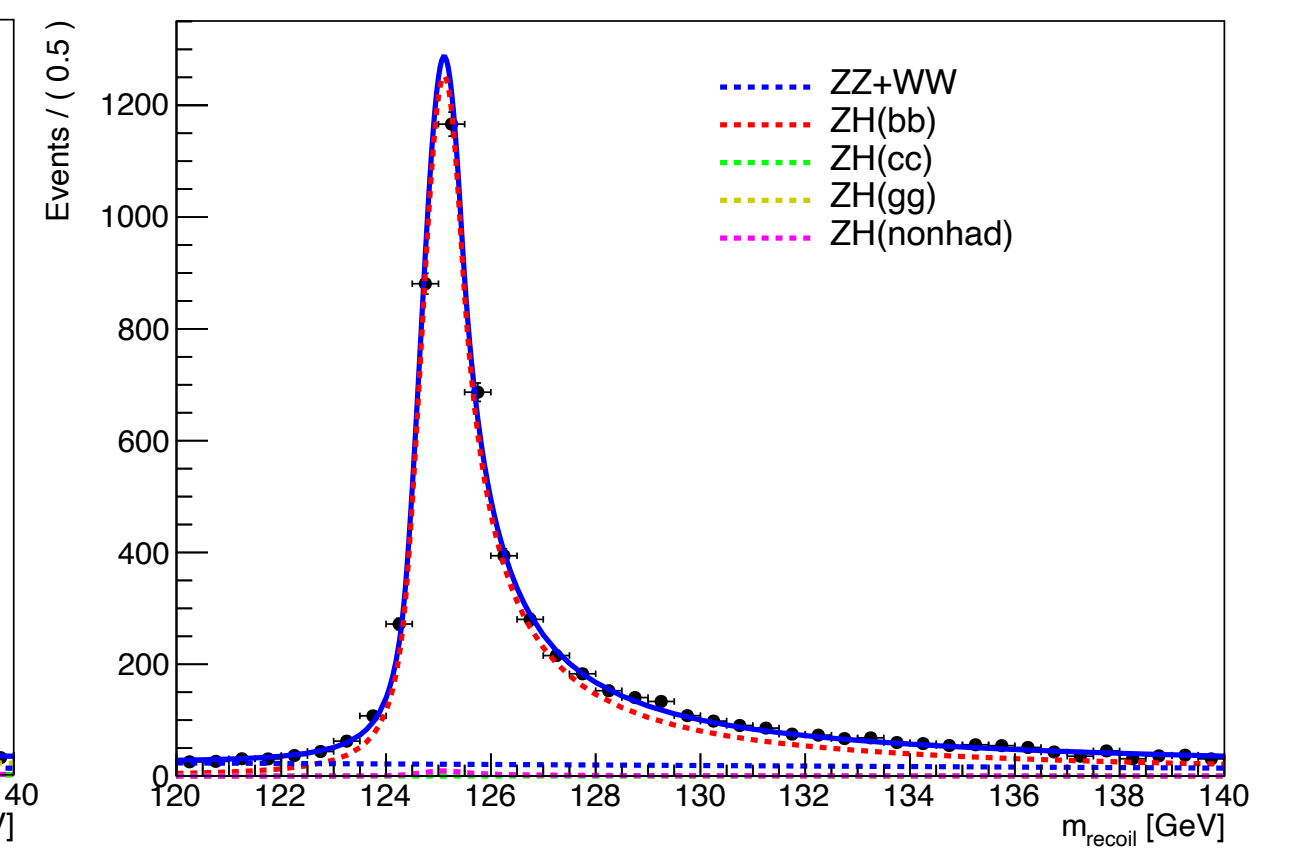
ee -> ZH, WW, ZZ 2c



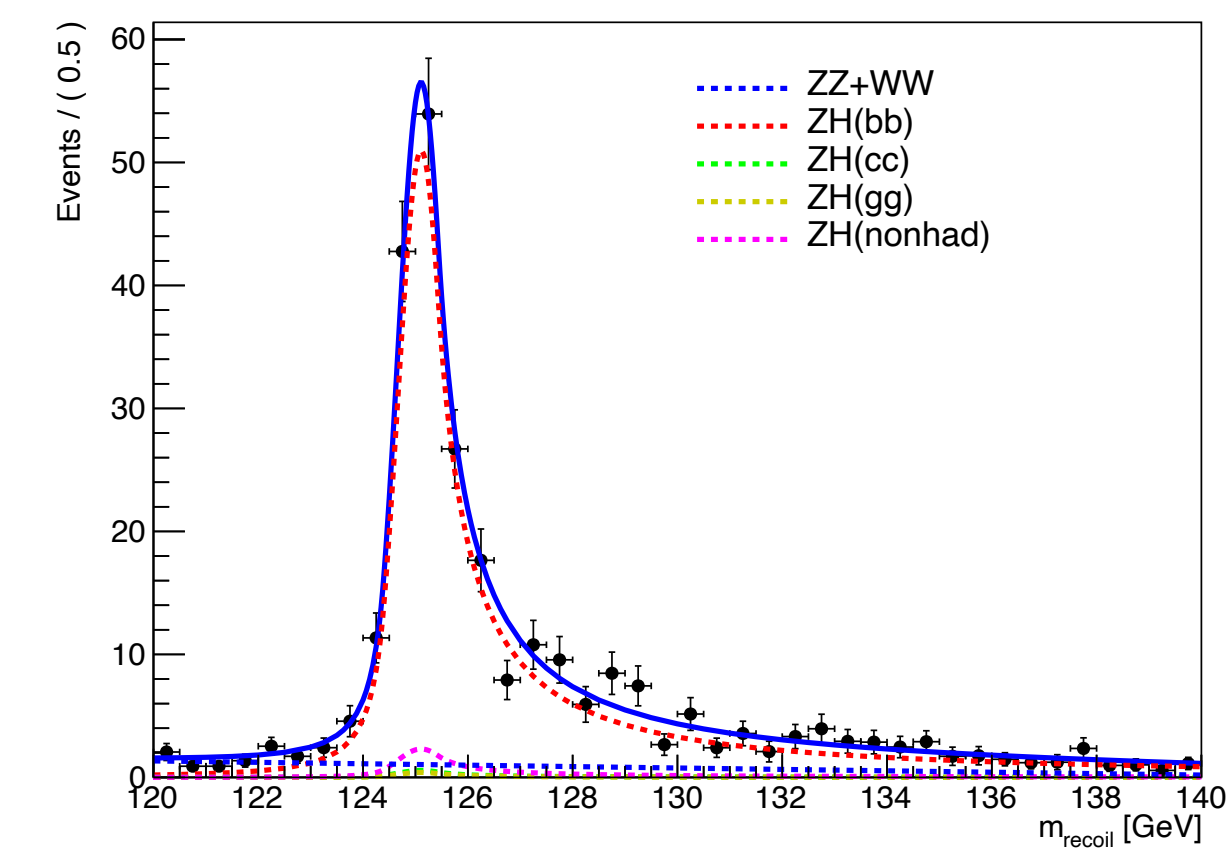
ee -> ZH, WW, ZZ 0b0c2g



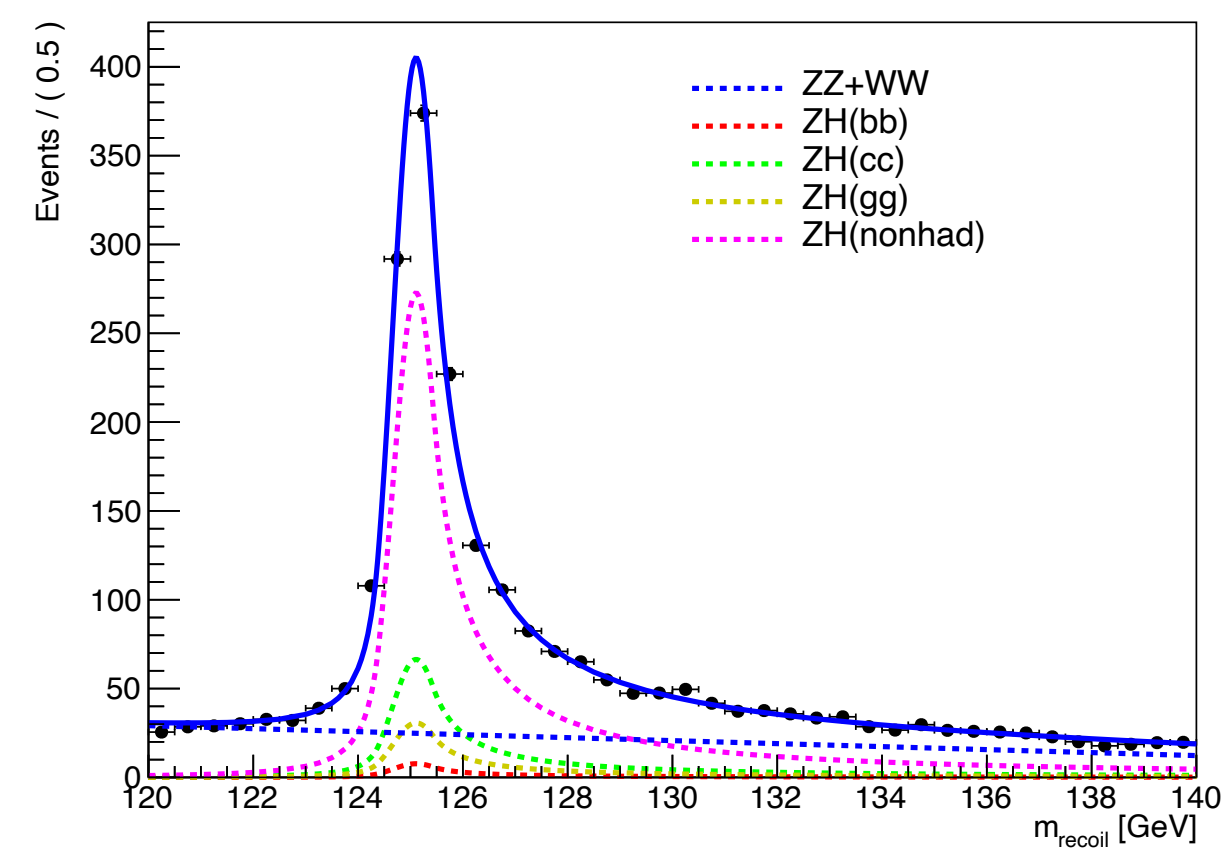
ee -> ZH, WW, ZZ 1b0c



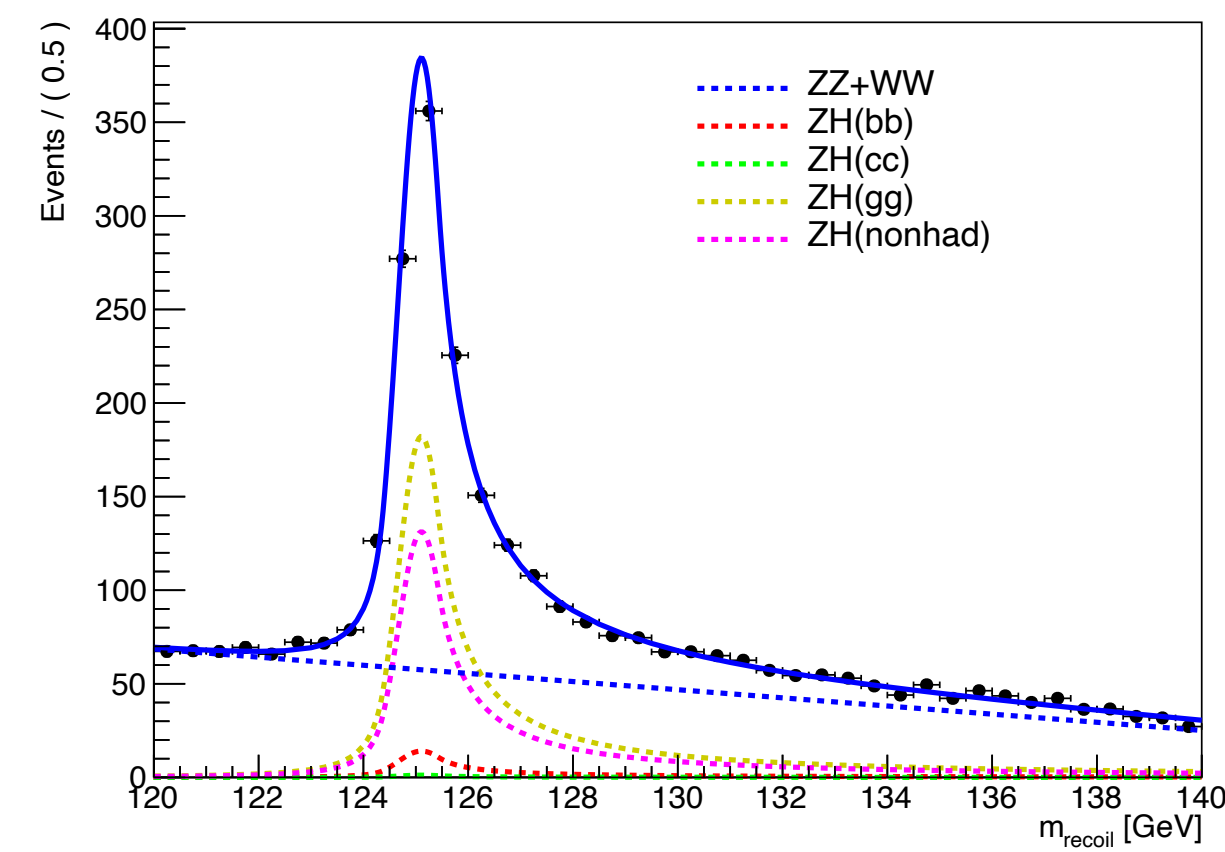
ee -> ZH, WW, ZZ 1b1c



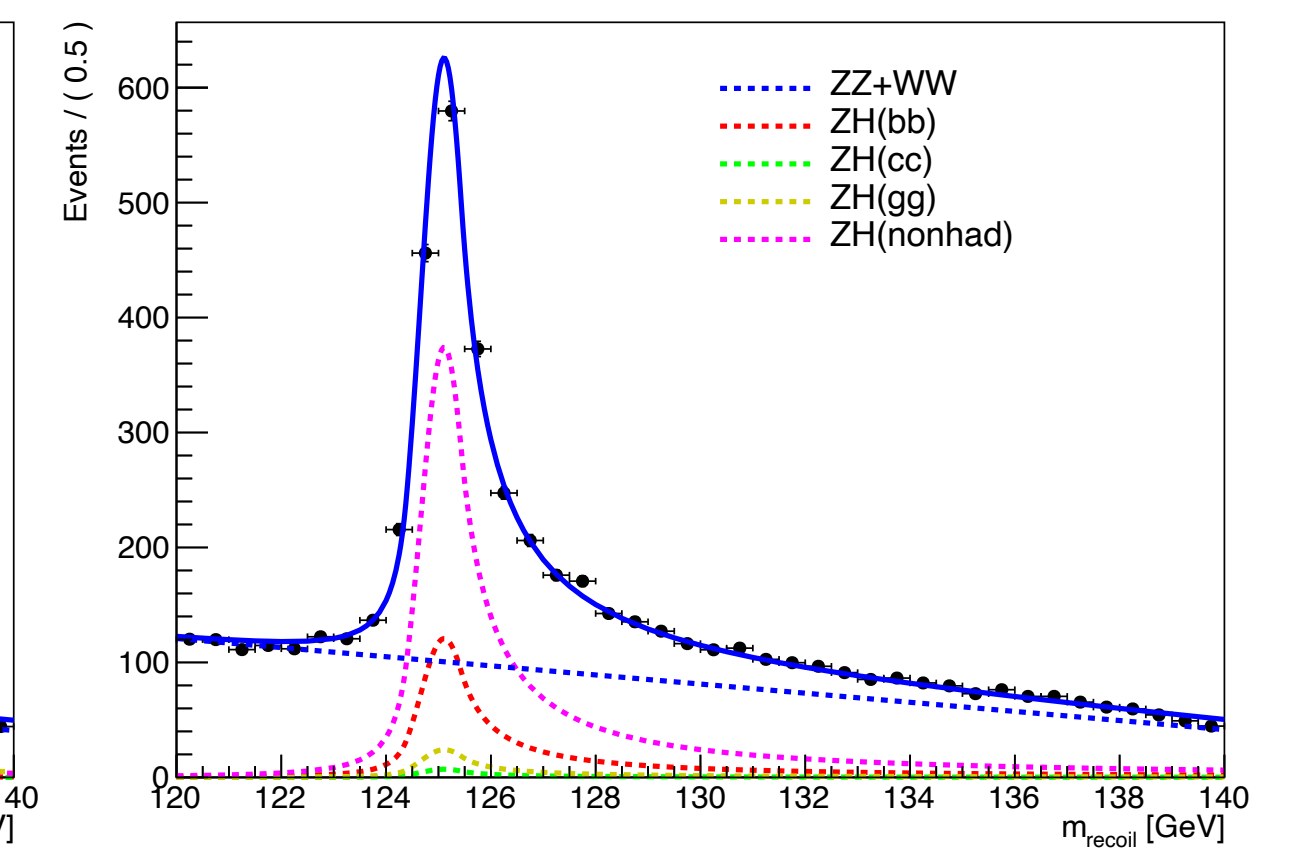
ee -> ZH, WW, ZZ 0b1c



ee -> ZH, WW, ZZ 0b0c1g



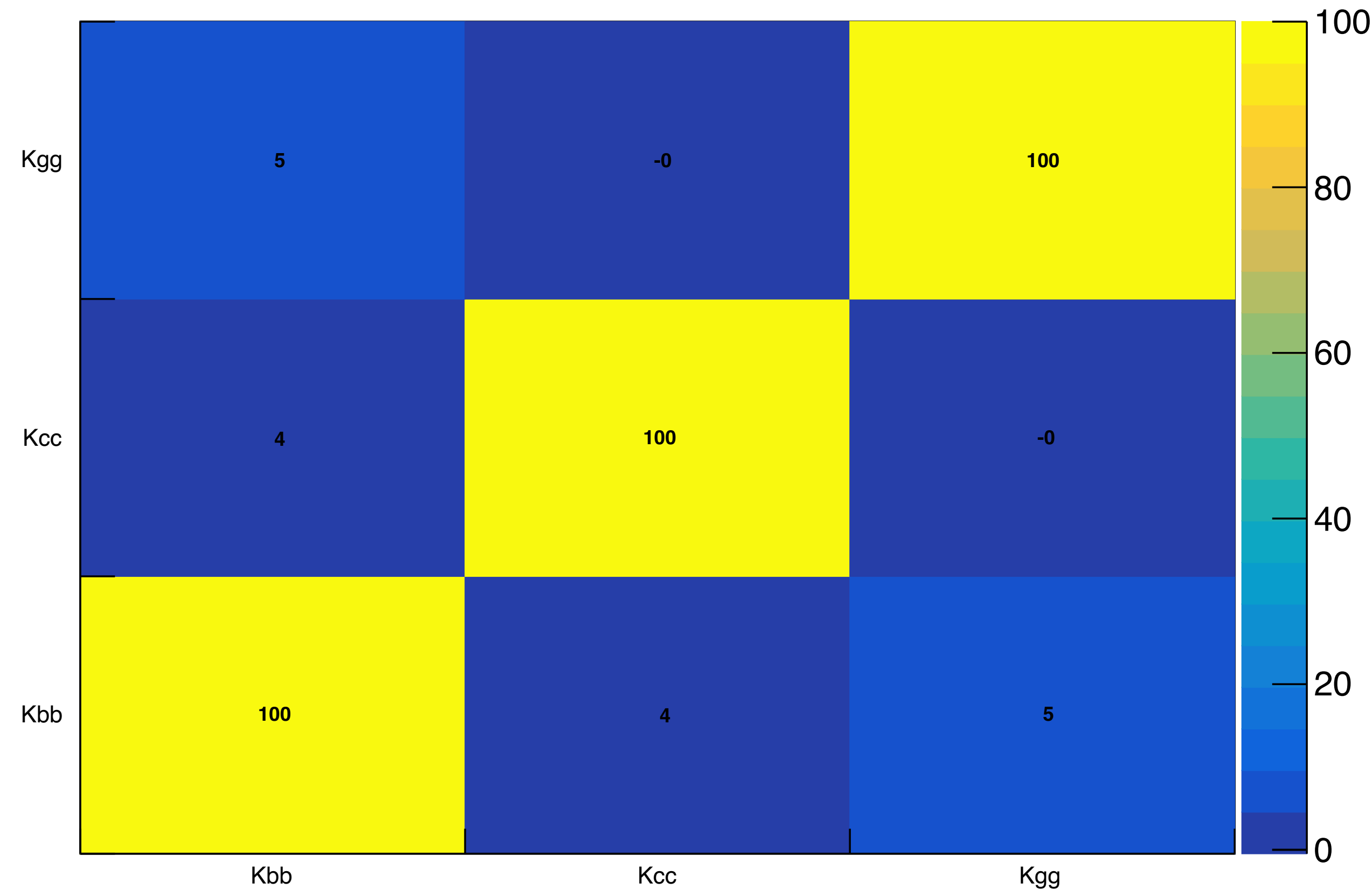
ee -> ZH, WW, ZZ 0b0c0g



Results - nominal selection

| Floating Parameter | InitialValue | FinalValue +/- | Error | GblCorr. |
|--------------------|--------------|----------------|----------|----------|
| Kbb | 1.0000e+00 | 1.0186e+00 +/- | 1.04e-02 | <none> |
| Kcc | 1.0000e+00 | 9.9378e-01 +/- | 6.45e-02 | <none> |
| Kgg | 1.0000e+00 | 1.0209e+00 +/- | 3.05e-02 | <none> |
| Yield_0b0c0g_bkg | 3.1857e+03 | 3.2476e+03 +/- | 6.79e+01 | <none> |
| Yield_0b0c1g_bkg | 1.8346e+03 | 1.8739e+03 +/- | 5.34e+01 | <none> |
| Yield_0b0c2g_bkg | 2.4100e+02 | 2.4876e+02 +/- | 3.28e+01 | <none> |
| Yield_0b1c_bkg | 7.9846e+02 | 8.2776e+02 +/- | 4.07e+01 | <none> |
| Yield_1b0c_bkg | 6.7222e+02 | 7.5469e+02 +/- | 5.66e+01 | <none> |
| Yield_1b1c_bkg | 3.3634e+01 | 3.0585e+01 +/- | 1.06e+01 | <none> |
| Yield_2b_bkg | 8.1358e+02 | 9.1335e+02 +/- | 8.16e+01 | <none> |
| Yield_2c_bkg | 7.9755e+02 | 8.2300e+02 +/- | 4.00e+01 | <none> |
| mH | 1.2500e+02 | 1.2499e+02 +/- | 5.79e-03 | <none> |
| p0_0b0c0g | 5.0000e+03 | 2.9092e+00 +/- | 2.41e-01 | <none> |
| p0_0b0c1g | 5.0000e+03 | 2.7340e+00 +/- | 2.91e-01 | <none> |
| p0_0b0c2g | 5.0000e+03 | 2.6256e+00 +/- | 1.10e+00 | <none> |
| p0_0b1c | 5.0000e+03 | 2.3736e+00 +/- | 4.07e-01 | <none> |
| p0_1b0c | 5.0000e+03 | 1.6573e+00 +/- | 3.49e-01 | <none> |
| p0_1b1c | 5.0000e+03 | 7.0104e+00 +/- | 1.67e+01 | <none> |
| p0_2b | 5.0000e+03 | 3.3001e+00 +/- | 1.01e+00 | <none> |
| p0_2c | 5.0000e+03 | 2.8467e+00 +/- | 4.65e-01 | <none> |
| sigma | 4.7100e-01 | 4.4326e-01 +/- | 7.95e-03 | <none> |

POI correlation matrix



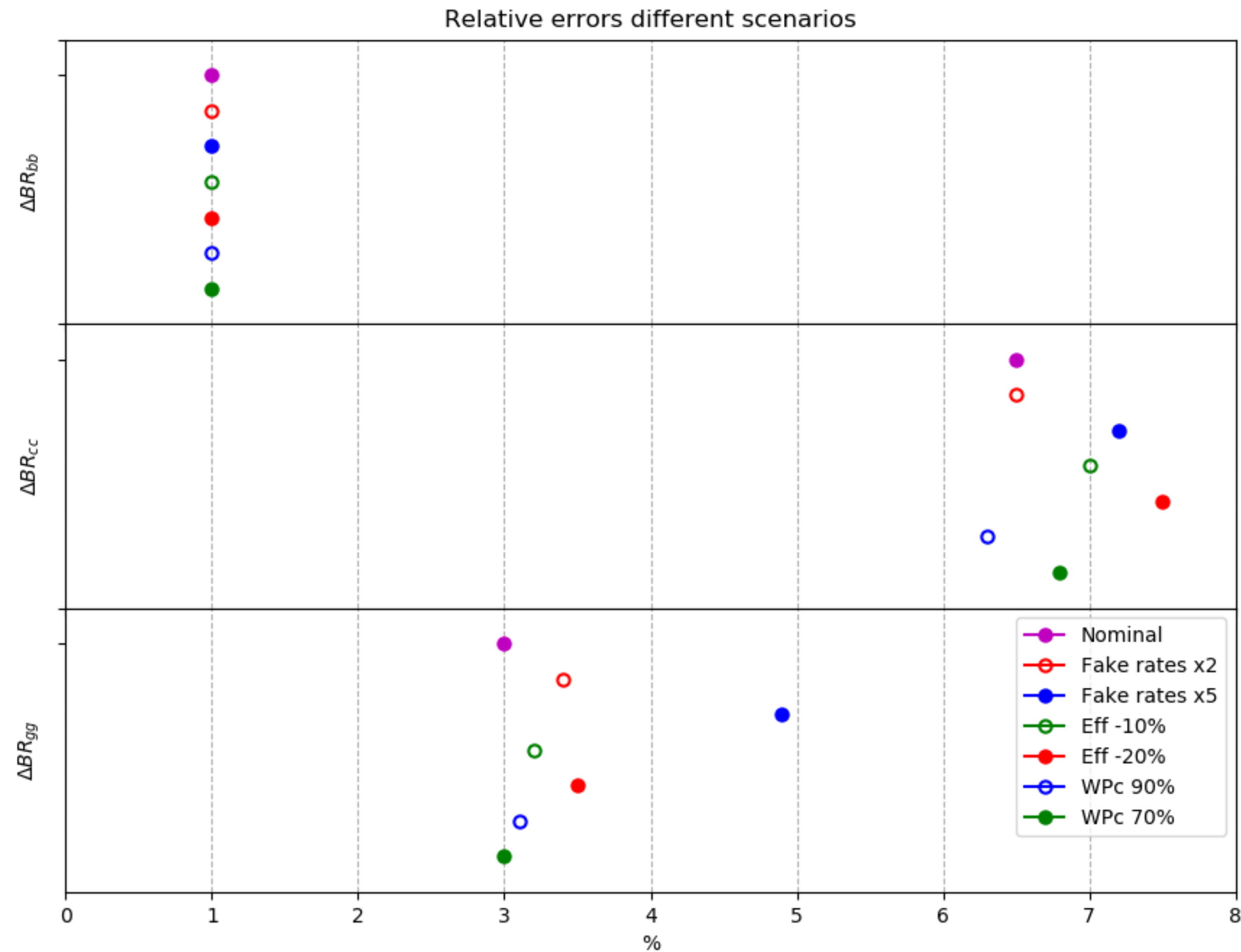
- Relative error on Kbb: 1.0%
- Relative error on Kcc: 6.5%
- Relative error on Kgg: 3.0% (3.1% when including also Z/y* bkg - not included in next slides)
- Relative error on Knonhad: 0.0% - fixed (3.4% when floating - leads to increase of Kcc uncertainty to 6.8%)

Results with various tagging performance/WP scenarios

| Strategy | ΔBR_{bb} (rel., %) | ΔBR_{cc} (rel., %) | ΔBR_{gg} (rel., %) |
|---------------|----------------------------|----------------------------|----------------------------|
| Nominal | 1.0 | 6.5 | 3.0 |
| Fake rates x2 | 1.0 | 6.5 | 3.4 |
| Fake rates x5 | 1.0 | 7.2 | 4.9 |
| Eff -10% | 1.0 | 7.0 | 3.2 |
| Eff -20% | 1.0 | 7.5 | 3.5 |
| WPc 90% | 1.0 | 6.3 | 3.1 |
| WPc 70% | 1.0 | 6.8 | 3.0 |

- Worse fake rate or efficiency do not affect significantly BR_{bb} thanks to large yield and purity, but can degrade BR_{cc} and BR_{gg} measurement - typically $O(15\%)$, with 50% for BR_{gg} in very extreme case of x5 $u \leftrightarrow g$ fake rate
- few% improvement on BR_{cc} (but slightly worse BR_{gg}) when going to higher-eff WP for c-tagging

Results with various tagging performance/WP scenarios



- Worse fake rate or efficiency do not affect significantly BR_{bb} thanks to large yield and purity, but can degrade BR_{cc} and BR_{gg} measurement - typically $O(15\%)$, with 50% for BR_{gg} in very extreme case of x5 $u \leftrightarrow g$ fake rate
- few% improvement on BR_{cc} (but slightly worse BR_{gg}) when going to higher-eff WP for c-tagging

Using non isolated leptons

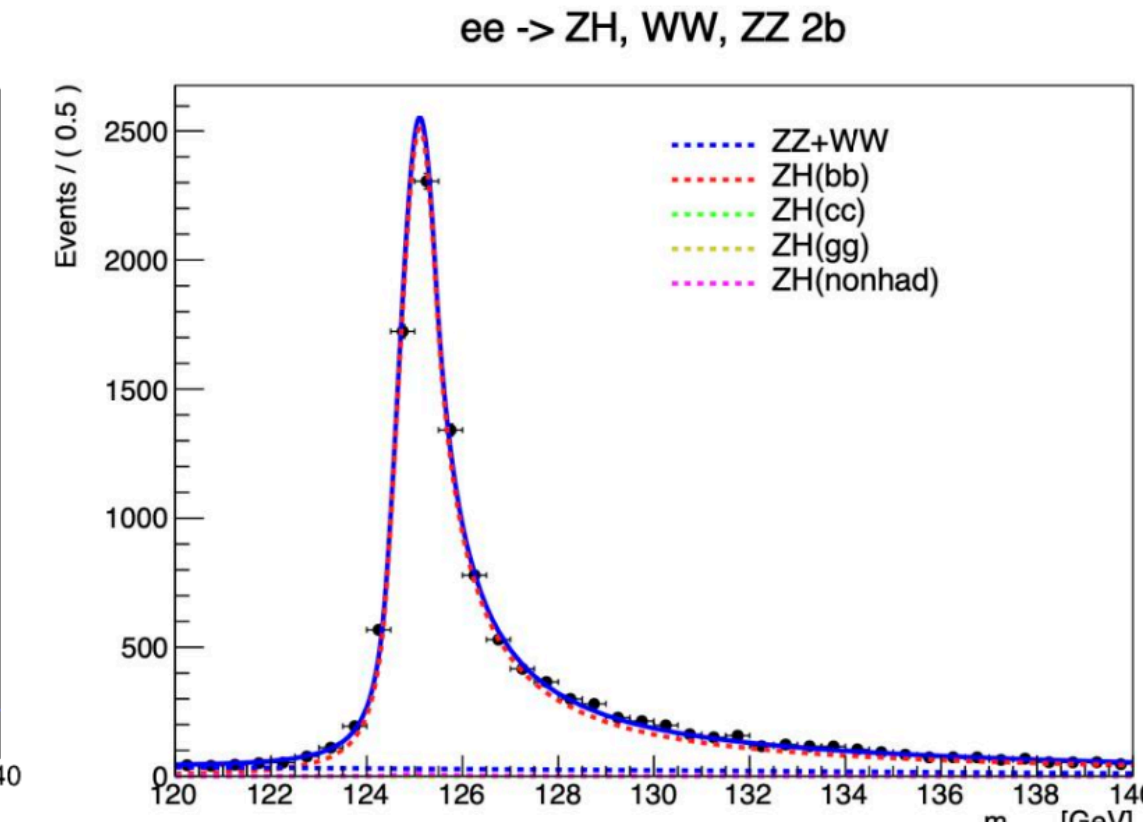
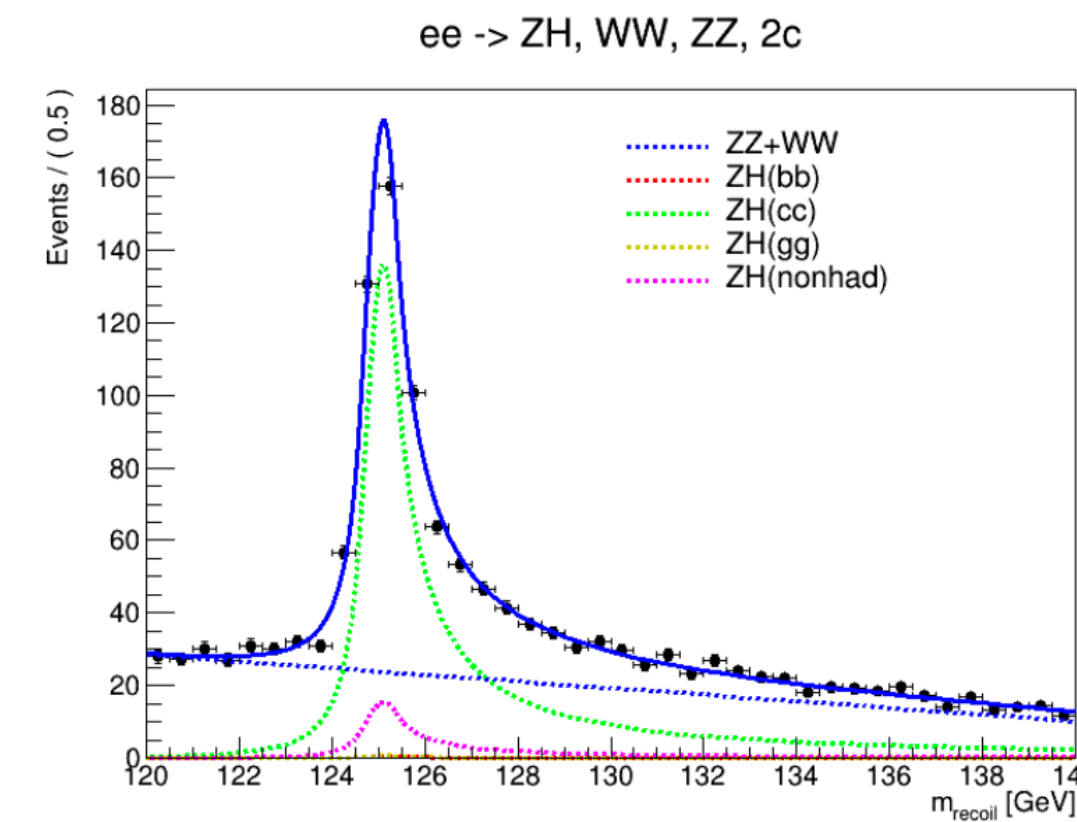
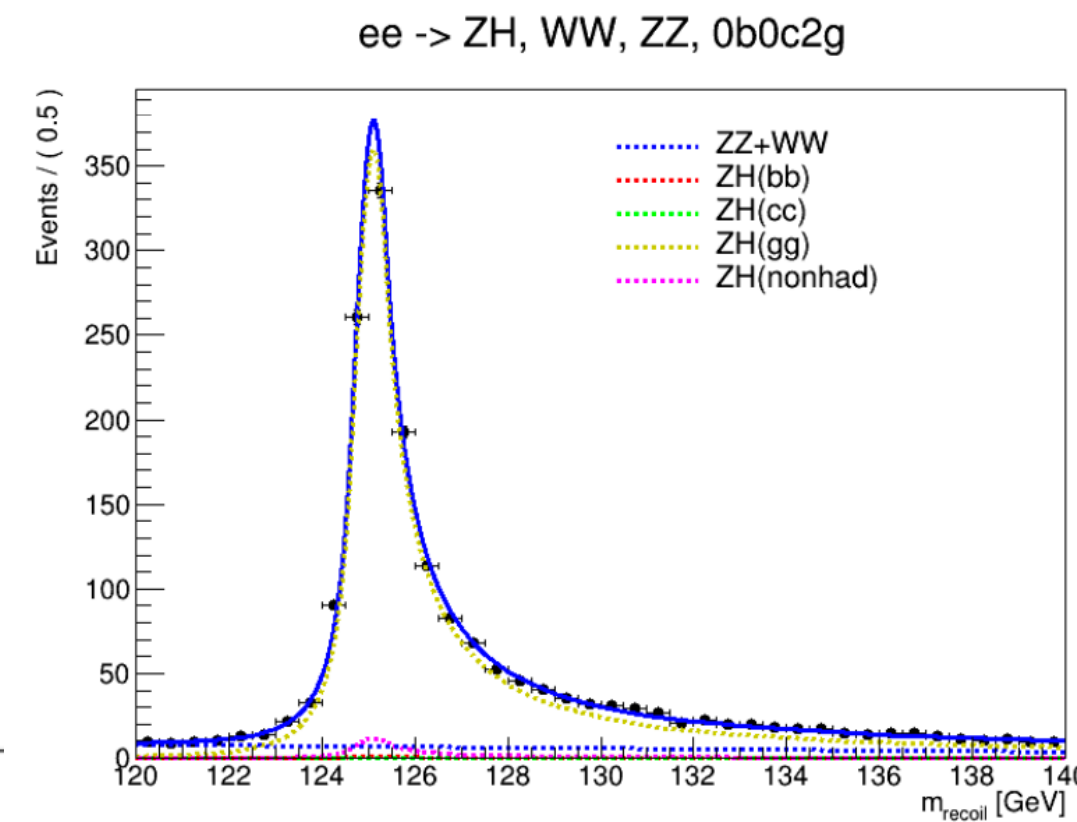
- Regenerating signals and main (WW, ZZ) bkg and saving non isolated electrons and leptons, we found an overall increase in efficiency of 30% for both signal and (main) bkg (~40% for ee and ~15% for mumu channels):

| Z(ee) | ZH _{bb} (%) | ZH _{cc} (%) | ZH _{gg} (%) | ZH _{other} (%) |
|--------------------|----------------------|----------------------|----------------------|-------------------------|
| Isolated leptons | 36.7 | 39.1 | 38.3 | 14.3 |
| No iso requirement | 51.3 | 55.3 | 54.5 | 22.4 |
| ratio | 1.40 | 1.41 | 1.42 | 1.57 |

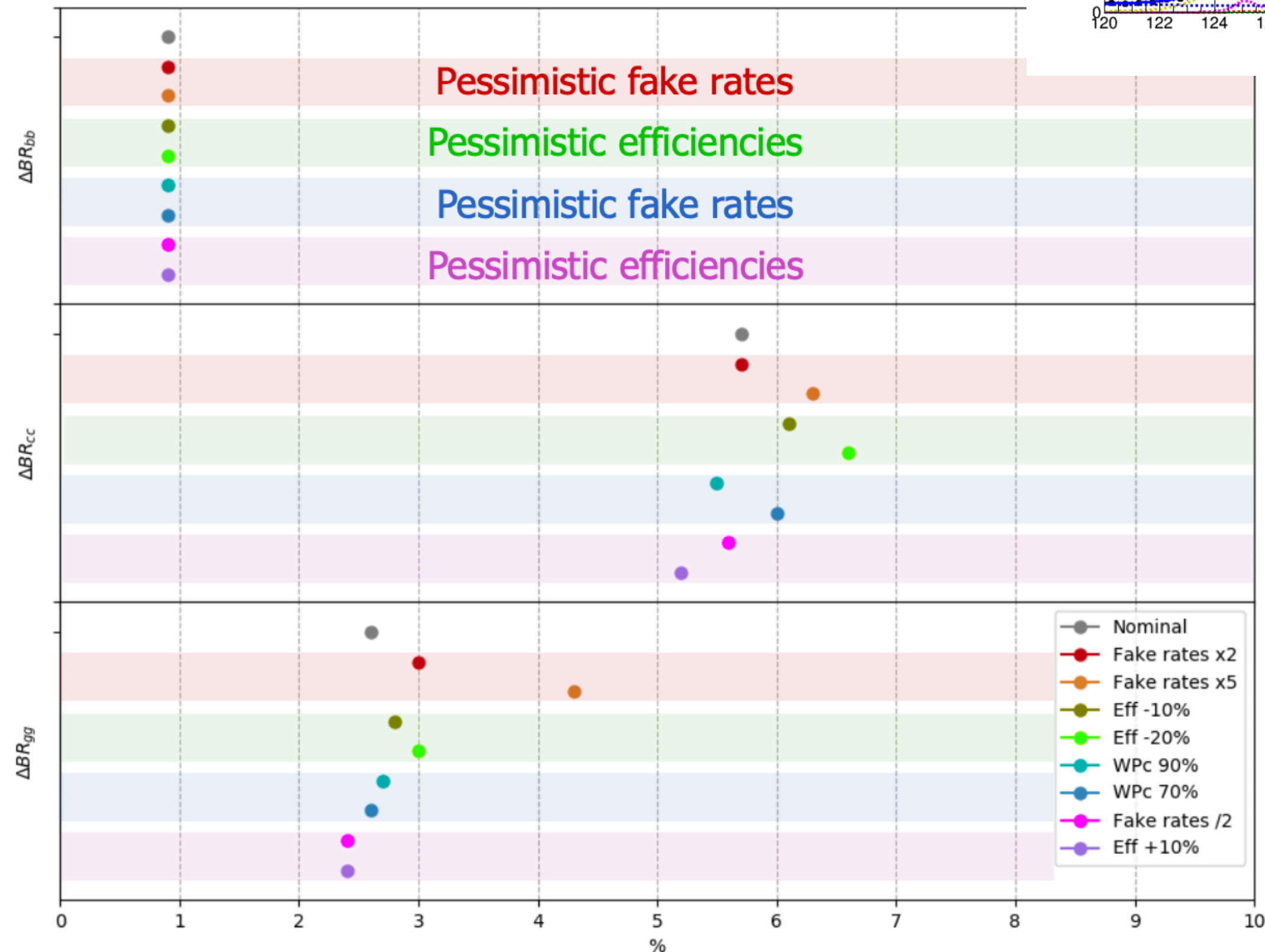
| Z(mumu) | ZH _{bb} (%) | ZH _{cc} (%) | ZH _{gg} (%) | ZH _{other} (%) |
|--------------------|----------------------|----------------------|----------------------|-------------------------|
| Isolated leptons | 48.0 | 51.5 | 52.4 | 19.7 |
| No iso requirement | 54.5 | 59.5 | 59.2 | 24.1 |
| ratio | 1.14 | 1.16 | 1.13 | 1.22 |

Using non isolated leptons

- Rescaling the signal and background by an overall factor 1.3 we arrive to new estimates of the expected yields and BR sensitivities:



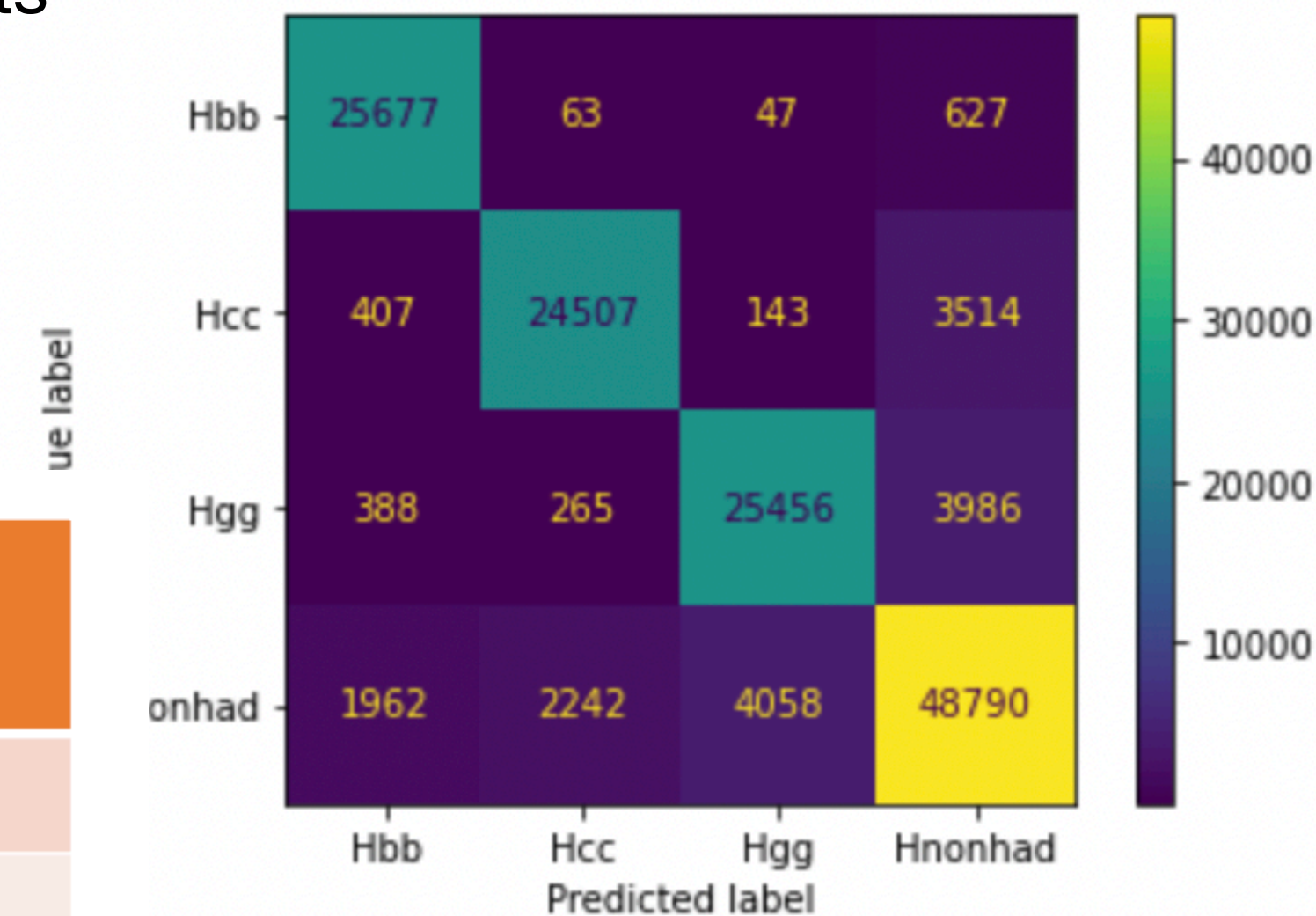
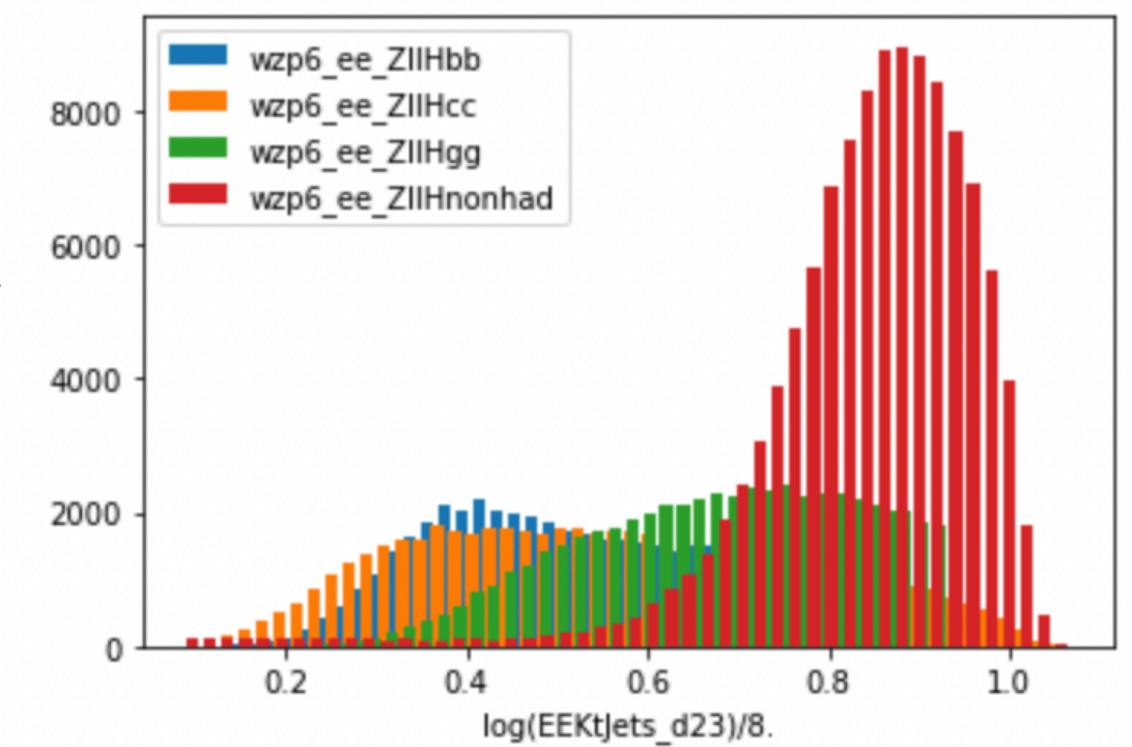
Relative errors different scenarios, lum = 5000*1.3



| Channel | Expected error |
|--------------------------|----------------|
| $H \rightarrow b\bar{b}$ | 0.9% |
| $H \rightarrow c\bar{c}$ | 5.9% |
| $H \rightarrow gg$ | 2.6% |

Further improvements

- Improve classification strategy: increase accuracy while simplifying number of categories
 - Use deep neural network based on number of b/c/g tags + few extra variables (Emiss, d_23 from last step of jet clustering, number of extra isolated leptons)
- Train multi-class NN with Z(ll) + H(bb)/H(cc)/H(gg)/H(other) as 4 training samples and targets
- Use output label of NN to classify events in 4 categories (rather than 8)
- Same fit strategy as before to determine bkg yields/shape parameters and signal $\sigma \cdot BR$
- *Preliminary* results show similar or even better sensitivity than previous strategy

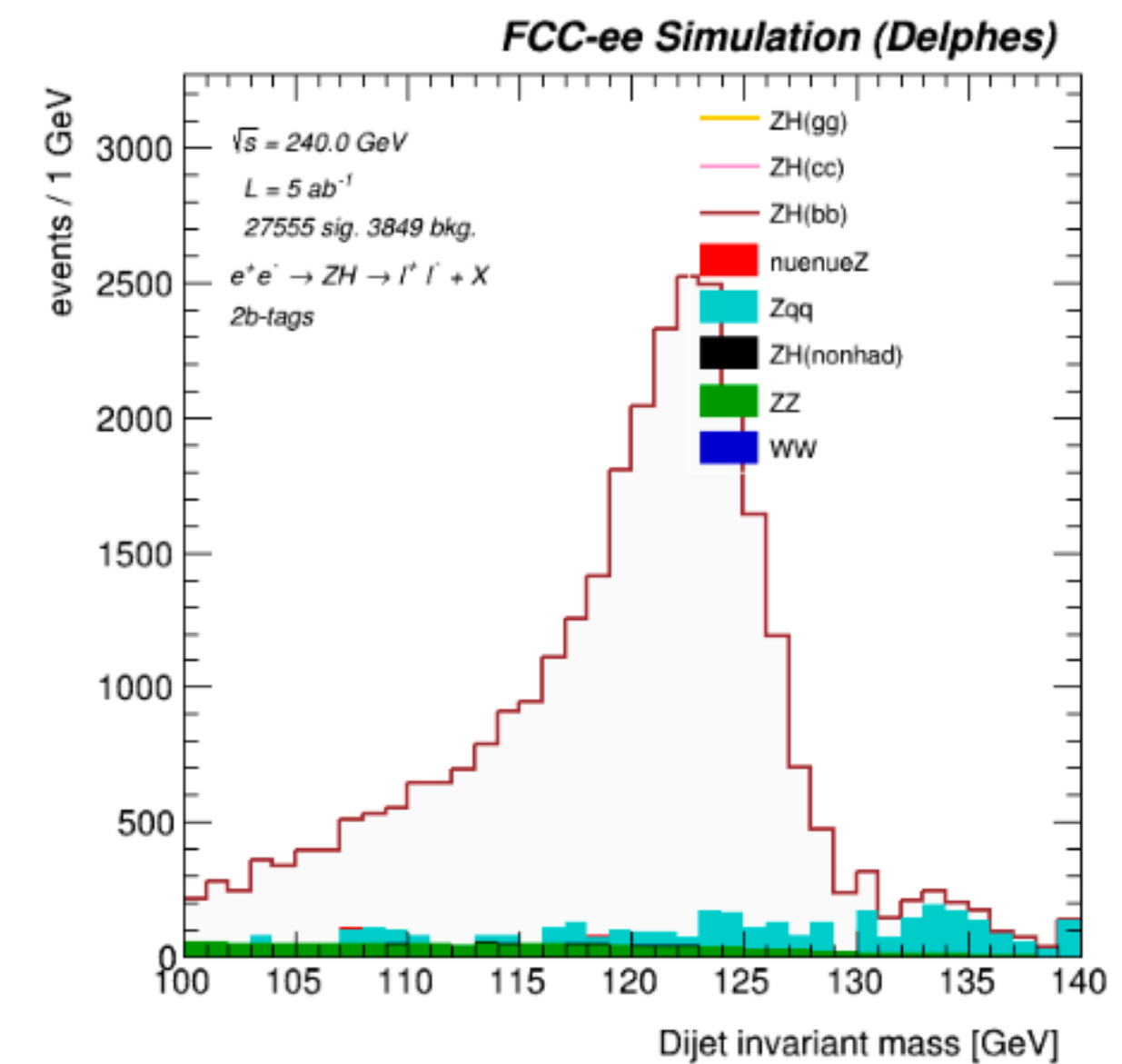
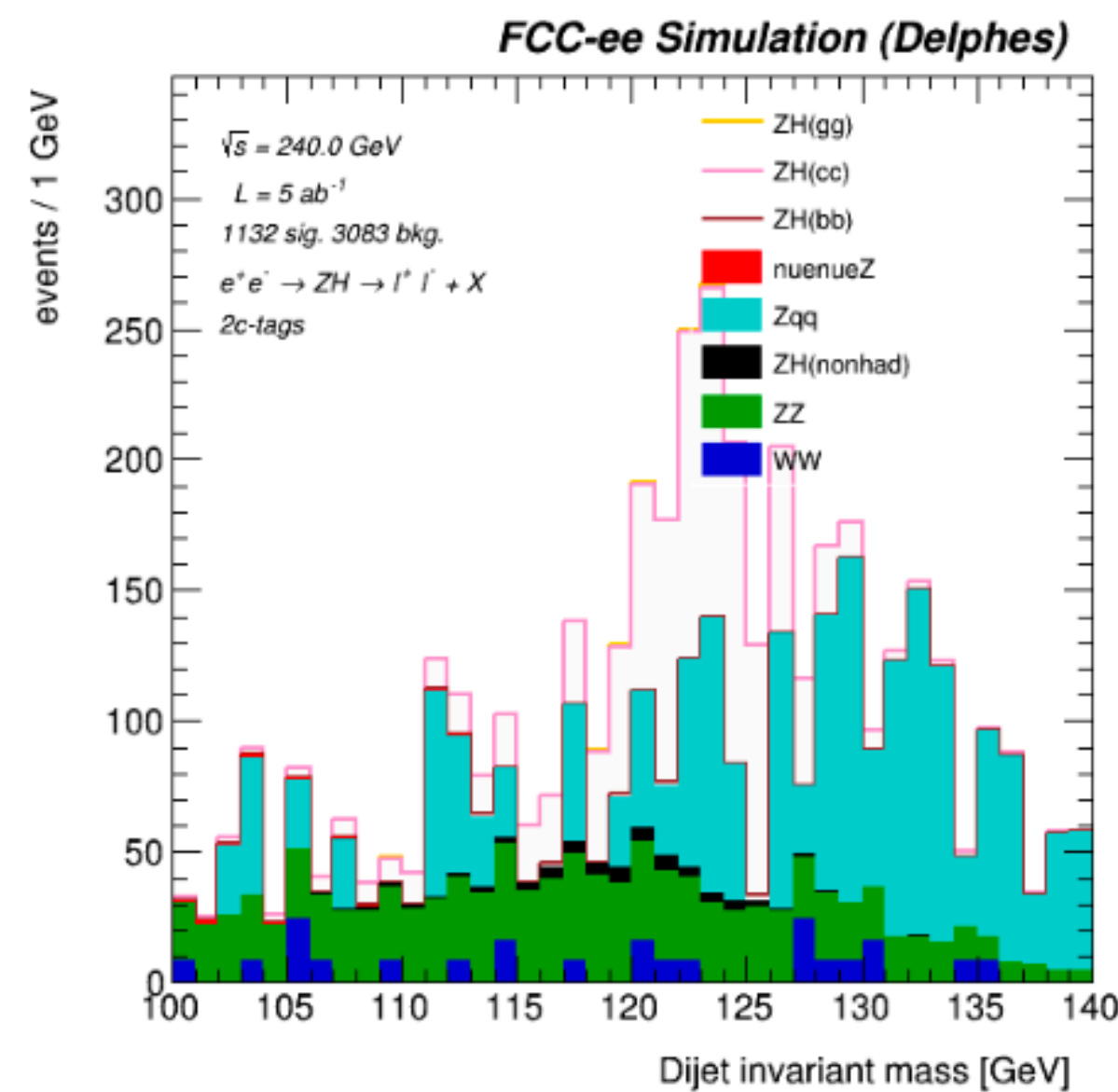
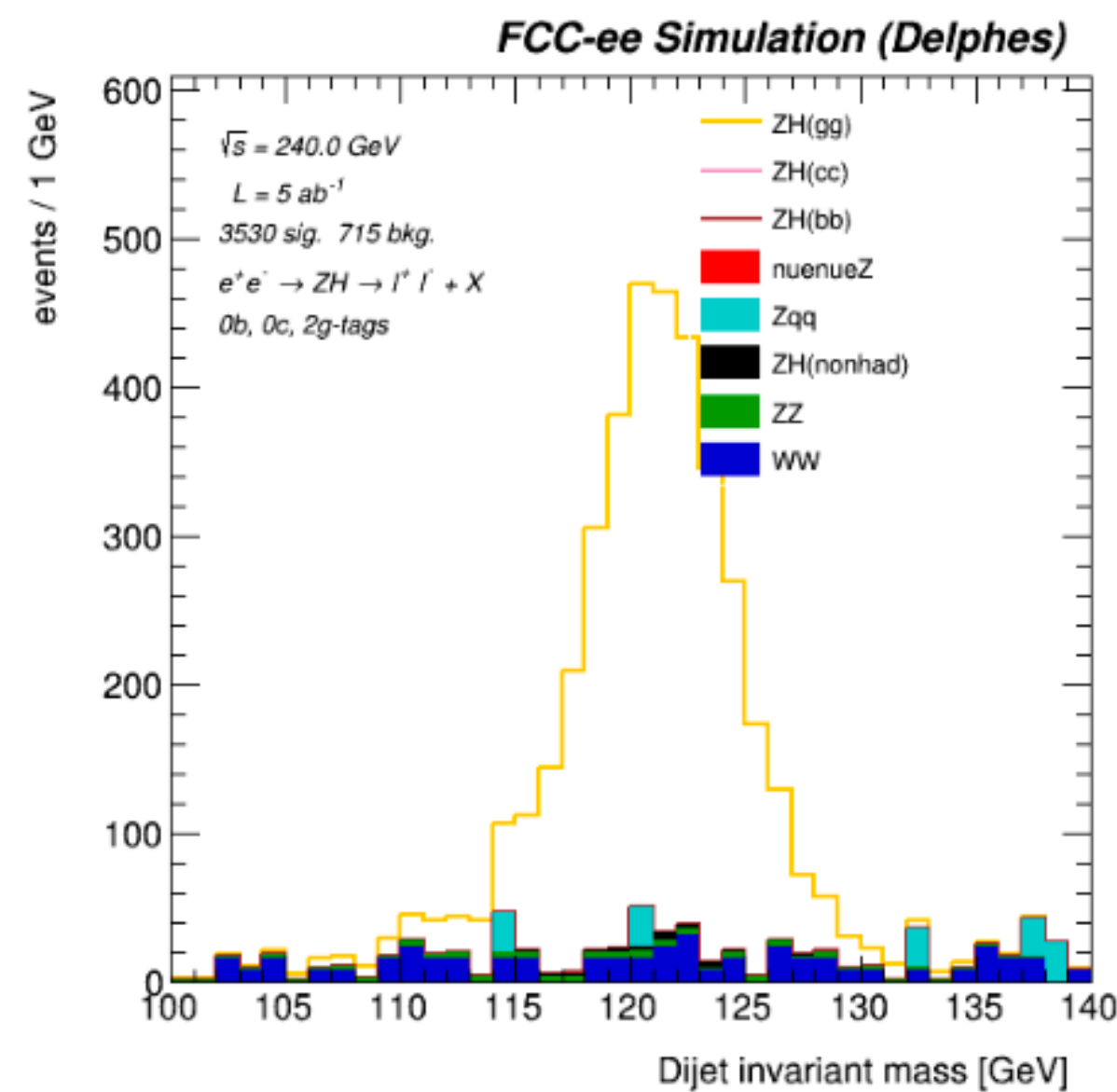


| Categories | Rel. Err. cc | Rel. Err. gg | Rel. Err. bb | Rel. Err. nonhad |
|-------------|--------------|--------------|--------------|------------------|
| 8 tags | 6.3 % | 2.8 % | 1 % | 3.1% |
| 4 tags (NN) | 4.8 % | 2.5 % | 0.8 % | 2.3 % |

- More input variables can be added; score of tagger, once available, can be used rather than number of tags; NN output score(s) could be fitted rather than m_recoil to extract the signal $\sigma \cdot BR$ (could be more separating than doing a m_recoil fit in NN categories)

Outlook on ongoing and future activities

- Ongoing:
 - Migrate analysis to latest FCCAnalysis version (~done)
 - Split ee/mumu categories (different S/B) (~done)
 - Include output of ParticleNet as input to NN (tagger found to perform worse than in Michele's studies, debugging with Michele, couple of bugs already found plus possibly also some differences induced by different clustering algorithm and details of event generation / parton shower in different samples used for training vs analysis)
 - Analysis of Z(nunu)H(qq) channel (selection/categorisation in place, working on implementation of fit procedure)



- Writing an internal supporting document to describe all these activities (started before summer break, ongoing)

Outlook on ongoing and future activities

- Ideas for the farther future:
 - Include also $H(ss)$ (and maybe also $H(\tau_{\text{had}} \tau_{\text{had}})$) in the analysis as target processes for which to measure $\sigma \cdot BR$
 - Look at $Z(\text{had})H(qq/gg)$
 - Look at $\sqrt{s}=365$ GeV
 - Fit NN output with templates rather than fit of m_{recoil} distribution with analytic models
 - Include in global combination with “measurements” of other decay channels

Backup

CEPC projections

- CEPCv1 (5.6/ab at 250 GeV, B=3.5T)

| Z decay mode | $H \rightarrow b\bar{b}$ | $H \rightarrow c\bar{c}$ | $H \rightarrow gg$ |
|------------------------------|--------------------------|--------------------------|--------------------|
| $Z \rightarrow e^+e^-$ | 1.3% | 12.8% | 6.8% |
| $Z \rightarrow \mu^+\mu^-$ | 1.0% | 9.4% | 4.9% |
| $Z \rightarrow q\bar{q}$ | 0.5% | 10.6% | 3.5% |
| $Z \rightarrow \nu\bar{\nu}$ | 0.4% | 3.7% | 1.4% |
| Combination | 0.3% | 3.1% | 1.2% |

- ~5% worse results for CEPCv4 vs v1

- CEPCv1 vs CEPCv4 (5.6/ab at 240 GeV, B=3T)

| Property | Estimated Precision | | | |
|-------------------------|---------------------|--|---------|--|
| | CEPC-v1 | | CEPC-v4 | |
| m_H | 5.9 MeV | | 5.9 MeV | |
| Γ_H | 2.7% | | 2.8% | |
| $\sigma(ZH)$ | 0.5% | | 0.5% | |
| $\sigma(\nu\bar{\nu}H)$ | 3.0% | | 3.2% | |

| Decay mode | $\sigma \times \text{BR}$ | BR | $\sigma \times \text{BR}$ | BR |
|---------------------------------------|---------------------------|---------|---------------------------|---------|
| $H \rightarrow b\bar{b}$ | 0.26% | 0.56% | 0.27% | 0.56% |
| $H \rightarrow c\bar{c}$ | 3.1% | 3.1% | 3.3% | 3.3% |
| $H \rightarrow gg$ | 1.2% | 1.3% | 1.3% | 1.4% |
| $H \rightarrow WW^*$ | 0.9% | 1.1% | 1.0% | 1.1% |
| $H \rightarrow ZZ^*$ | 4.9% | 5.0% | 5.1% | 5.1% |
| $H \rightarrow \gamma\gamma$ | 6.2% | 6.2% | 6.8% | 6.9% |
| $H \rightarrow Z\gamma$ | 13% | 13% | 16% | 16% |
| $H \rightarrow \tau^+\tau^-$ | 0.8% | 0.9% | 0.8% | 1.0% |
| $H \rightarrow \mu^+\mu^-$ | 16% | 16% | 17% | 17% |
| $\text{BR}_{\text{inv}}^{\text{BSM}}$ | — | < 0.28% | — | < 0.30% |

FCCee CDR projections

Table 4.1 Relative statistical uncertainty on the measurements of event rates, providing $\sigma_{\text{HZ}} \times \text{BR}(\text{H} \rightarrow \text{XX})$ and $\sigma_{\nu\bar{\nu}\text{H}} \times \text{BR}(\text{H} \rightarrow \text{XX})$, as expected from the FCC-ee data. This is obtained from a fast simulation of the CLD detector and consolidated with extrapolations from full simulations of similar linear-collider detectors (SiD and CLIC). All numbers indicate 68% C.L. intervals, except for the 95% C.L. sensitivity in the last line. The accuracies expected with 5 ab^{-1} at 240 GeV are given in the middle columns, and those expected with 1.5 ab^{-1} at $\sqrt{s} = 365 \text{ GeV}$ are displayed in the last columns

| \sqrt{s} (GeV) | 240 | | 365 | |
|---|-----------|------------------------|-----------|------------------------|
| Luminosity (ab^{-1}) | 5 | | 1.5 | |
| $\delta(\sigma\text{BR})/\sigma\text{BR}$ (%) | HZ | $\nu\bar{\nu}\text{H}$ | HZ | $\nu\bar{\nu}\text{H}$ |
| H \rightarrow any | ± 0.5 | | ± 0.9 | |
| H \rightarrow $b\bar{b}$ | ± 0.3 | ± 3.1 | ± 0.5 | ± 0.9 |
| H \rightarrow $c\bar{c}$ | ± 2.2 | | ± 6.5 | ± 10 |
| H \rightarrow gg | ± 1.9 | | ± 3.5 | ± 4.5 |
| H \rightarrow W^+W^- | ± 1.2 | | ± 2.6 | ± 3.0 |
| H \rightarrow ZZ | ± 4.4 | | ± 12 | ± 10 |
| H \rightarrow $\tau\tau$ | ± 0.9 | | ± 1.8 | ± 8 |
| H \rightarrow $\gamma\gamma$ | ± 9.0 | | ± 18 | ± 22 |
| H \rightarrow $\mu^+\mu^-$ | ± 19 | | ± 40 | |
| H \rightarrow invis. | < 0.3 | | < 0.6 | |

Projections in FCC CDR

- Assuming **5/ab at 240 GeV**, the projected sensitivities at FCC-ee in the CDR ([Eur. Phys. J. C 79, 474 \(2019\), Table 4.1](#)) scaled to the Z(ll) channel alone (taking Z(ll)H/Z(ll+qq+vv)H from Table 6 of CEPC, <https://arxiv.org/pdf/1810.09037.pdf>):

| Higgs decay mode | Rel. unc. on $\sigma\text{BR}(\text{bb})$ (%) |
|------------------|---|
| bb | 0.8 |
| cc | 5.3 |
| gg | 6.1 |

* Considering only background from ZZ and WW (background from ZH with other Higgs decays not considered)