$\mathrm{H} \to \mathrm{ZZ}^*$ @ 350 GeV and 3 TeV CLIC

Continuation of discussion on CLICdp-Draft-2021-010

CLICdp-Draft-2021-010, submitted on the 24th of June 2021

Natasa Vukasinovic, et al.
VINCA Belgrade & Uni. Kragujevac

Slides: 3 - 15

Follow-up material from the AWG meetings on the 19th of April and 17th of May

From slide 16:

Answers to questions raised after the 19th of July

Comments/questions to the analysis raised before the 19th of April:

1) Implement lepton dressing @ 350 GeV and 3 TeV

2) IP issue

   a. Relax e^- IP cuts @ 350 GeV (P. Roloff)

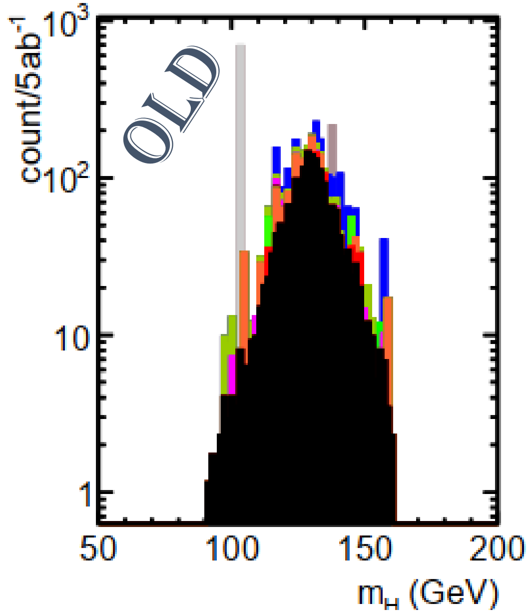
3) Background processes @ 3 TeV

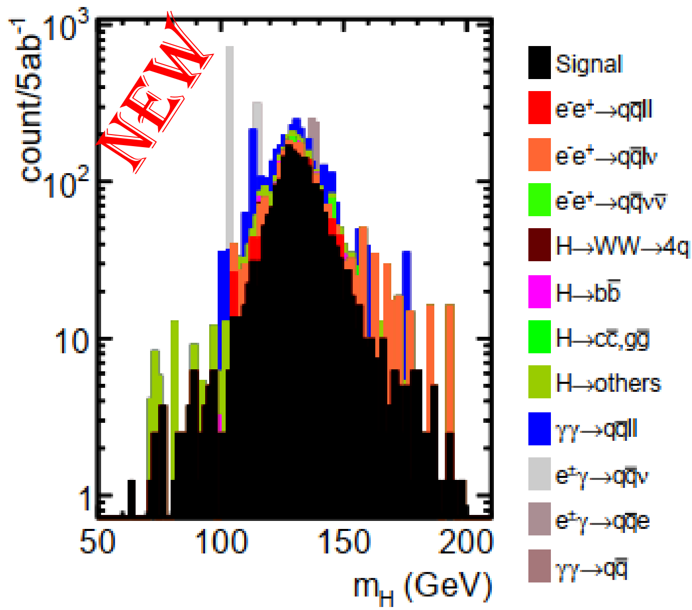
   a. Do you understand where the "cut" at ~130 GeV, visible both in signal and background distributions at 350 GeV, come froms (F. Zarnecki)?

   b. Why 3 TeV analysis does not include 6 fermion backgrounds. Have they been considered (F. Zarnecki)?

1) ANSWERED on the 19th of April AWG meeting

1) Impact of SIGNAL lepton dressing on MVA performance @ 3 TeV





OLD

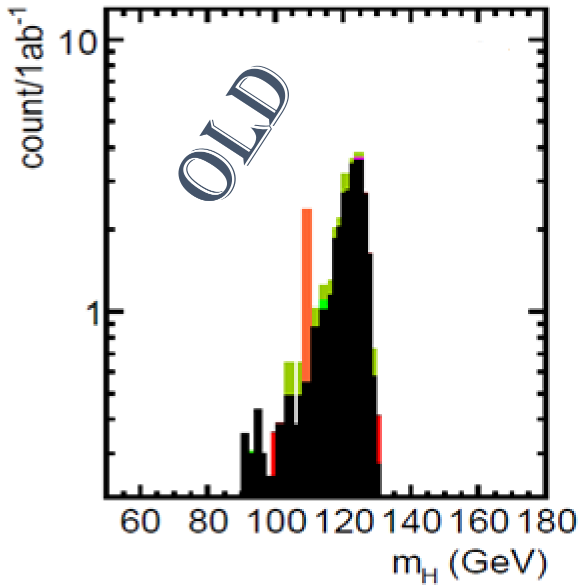
Signal @ 3 TeV	δ=3.4%	67%	47%	1793
Background process				
$e^{\pm}\gamma \to q\bar{q}\nu$		1.4‰	3.5‰	678
$\gamma\gamma \to q\bar{q}l^+l^-$		11‰	0.3‰	444
$e^-e^+ \to q\bar{q}l\nu$		3‰	3.4‰	281
$e^-e^+ \to H\nu\bar{\nu}; H \to others$		45‰	1%	211
$e^{\pm}\gamma \to q\bar{q}e$		8.8‰	0.2‰	116
processes with N_{BDT}<100		5.2‰	1.2‰	152

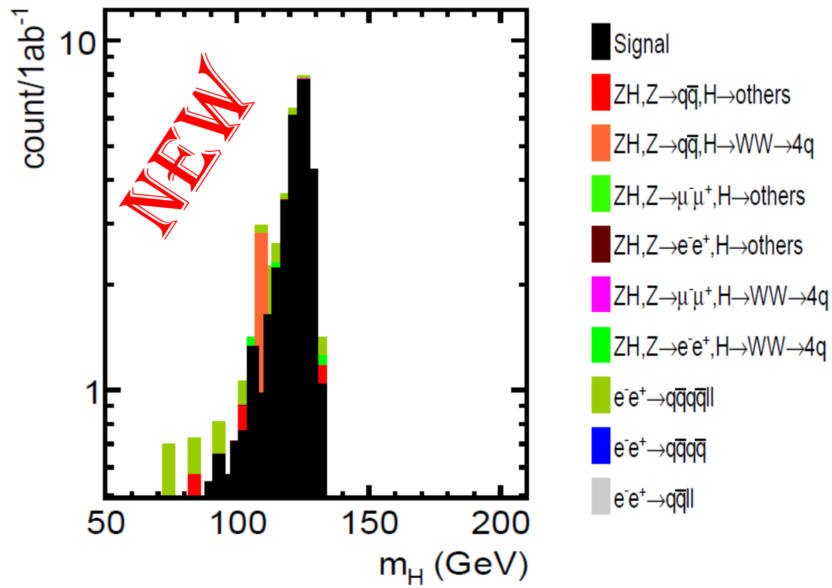
NEW

Signal @ 3 TeV	δ=3.2%	67%	60%	2281
Background process				
$e^{\pm}\gamma \to q\bar{q}\nu$		1.4‰	5‰	930
$\gamma\gamma \to q\bar{q}l^+l^-$		11‰	0.5‰	820
$e^-e^+ \to q\bar{q}l\nu$		3‰	6‰	482
$e^-e^+ \to H\nu\bar{\nu}; H \to others$		45‰	2%	345
$e^{\pm}\gamma \to q\bar{q}e$		8.8‰	0.4‰	232
processes with N_{BDT}<100		5.2‰	1.3‰	170

1) ANSWERED on the 19th of April AWG meeting

1) Impact of SIGNAL lepton dressing on MVA performance @ 350 GeV





δ=19% (OLD)

Process	ε_{presel}	ε_{BDT}	N_{BDT}
Signal @ 350 GeV	76%	17%	32
Background process			
$e^-e^+ \to q\bar{q}q\bar{q}l^+l^-$	21%	3‰	3
$e^-e^+ \to HZ;Z \to q\bar{q},H \to WW \to 4q$	0.42%	4.5%	2

δ=18.2% (NEW)

Process	ε_{presel}	ε_{BDT}	N_{BDT}
Signal @ 350 GeV	77%	19%	35
Background process			
$e^-e^+ \to q\bar{q}q\bar{q}l^+l^-$	21%	3‰	3
$e^-e^+ \to HZ;Z \to q\bar{q},H \to WW \to 4q$	0.42%	4.5%	2

2) ANSWERED on the 19th of April AWG meeting

2) IP issue - relaxed e^- IP cuts

- We've repeated the analysis at 350 GeV with d_0-cut of 10 mm and z_0- cut of 5 mm for electrons.

- Relaxing electron IP cuts increased the signal events but also the background in a way that MVA **significance dropped from 5.5 to 2.8, that is statistical error is 36% instead of previous 18.2%.**

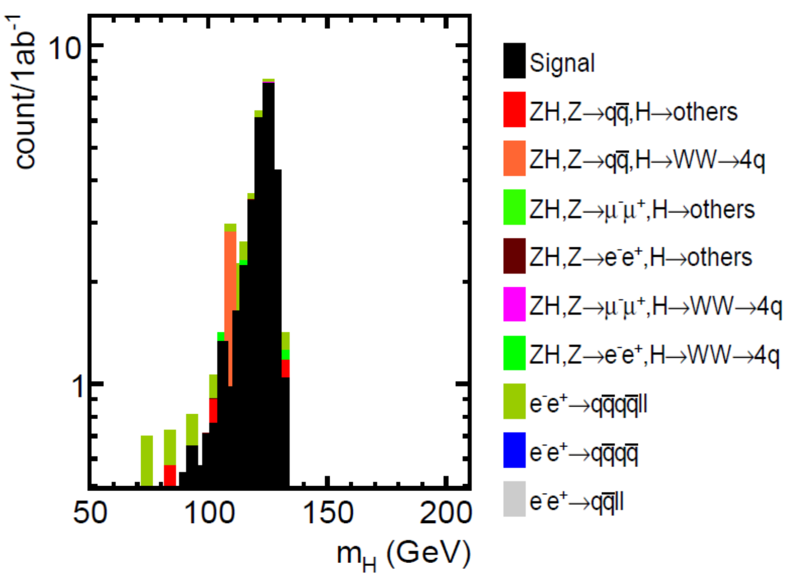
2) ANSWERED on the 19th of April AWG meeting

2) IP issue - relaxed e^- IP cuts

d_0=0.02mm; z_0=0.02mm

δ=18%

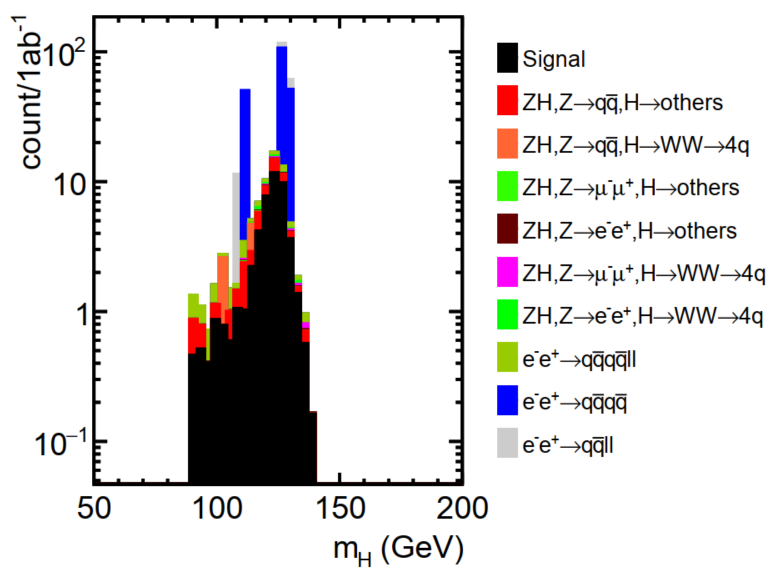
Process	ε_{presel}	ε_{BDT}	N_{BDT}
Signal @ 350 GeV	77%	19%	35
Background process			
$e^-e^+ \to q\bar{q}q\bar{q}l^+l^-$	21%	3‰	3
$e^-e^+ \to HZ;Z \to q\bar{q}, H \to WW \to 4q$	0.42%	4.5%	2



d_0=10mm; z_0=5mm

δ=36%

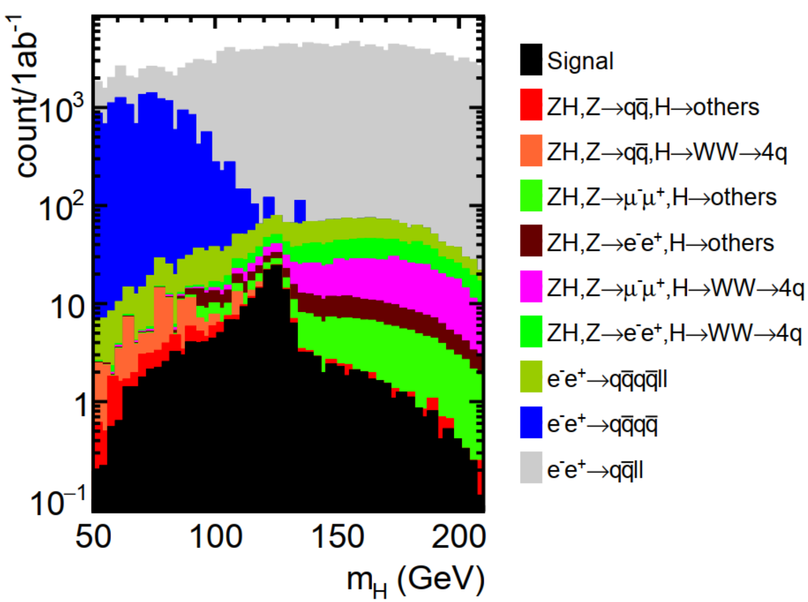
Process	ε_{presel}	ε_{BDT}	N_{BDT}
Signal @ 350 GeV	85%	23%	47
Background process			
$e^-e^+ \to q\bar{q}q\bar{q}$	0.32%	0.01%	187
$e^-e^+ \to q\bar{q}l^+l^-$	11.4%	0.1‰	29
$e^-e^+ \to HZ;Z \to q\bar{q}, H \to others$	0.37%	0.5%	12
$e^-e^+ \to q\bar{q}q\bar{q}l^+l^-$	21%	8‰	8
$e^-e^+ \to HZ;Z \to q\bar{q}, H \to WW \to q\bar{q}q\bar{q}$	0.42%	9%	4

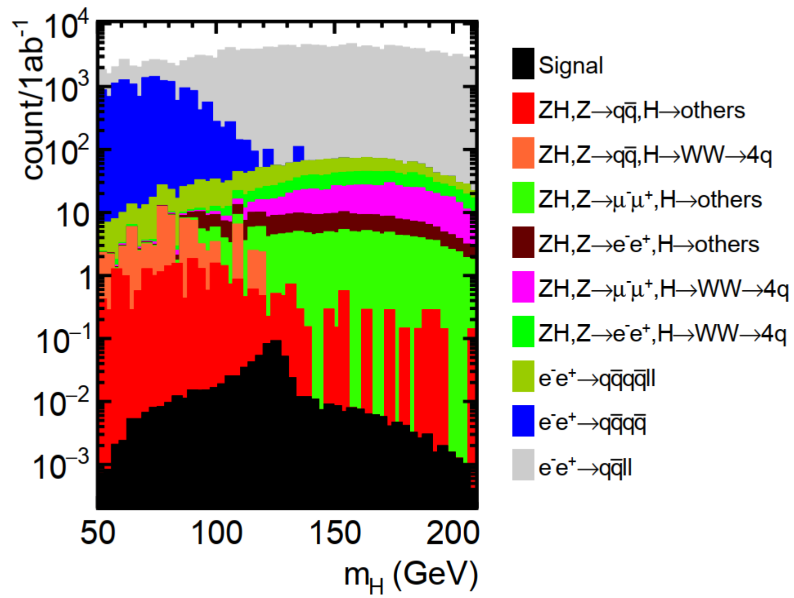


3a) ANSWERED on the 19th of April AWG meeting

Q: Do you understand where the "cut" at ~130 GeV, visible both in signal and background distributions at 350 GeV, come from?

- If we artificially rescale signal (multiply with a small factor (0.0001) – the shape stays, of course, unchanged) we see that only 3 processes (not all) die around 130 GeV (blue, orange and red).
- Two are multi-jet processes where primary Z decays hadronically. Orange is even a process with 6 jets. We would say that it naturally dies due to the lack of the phase space(energy) at 350 GeV.
- It seems to be the issue of presentation.





3b) ANSWERED on the 19th of April AWG meeting

3b) Background processes @ 3 TeV - qqqqll

Q: The table for 3TeV does not include 6 fermion backgrounds. Have they been considered?

- **Two processes have ben emphasized:**
  - **qqqqll** (Q: 'It seems strange to include eg. qqqqll at 350 GeV but not at 3 TeV?')
  - **qqllvv at 3 TeV**

A:

- Signal at 3 TeV has missing energy signature, differently from qqqqll events (and the signal in HZ)
- In our analysis of H to ZZ decays at 1.4 TeV (contribution to the Higgs paper), it is found that qqqqll can be fully suppressed with E_{miss} and jet-transition variables for the qqll final state. It can eventually play a role in the hadronic final state Hvv, H→ZZ→qqqq, where **it contributes with ~0.001% after MVA.**

Background	ε_{BDT} (%)	N_{BDT}
$e^+e^- \to q\bar{q}q\bar{q}\nu_e\bar{\nu}_e$	43.0	191
$e^+e^- \to q\bar{q}q\bar{q}l\nu$	0.1	48
$e^+e^- \to q\bar{q}q\bar{q}l^+l^-$	0.001	6
$e^+e^- \to q\bar{q}q\bar{q}$	0.008	6
$e^+e^- \to q\bar{q}$	100	0
$e^\pm\gamma \to q\bar{q}q\bar{q}e$	100	0
$\gamma\gamma \to q\bar{q}q\bar{q}$	100	0

CERN-THESIS-2017-349

3b) ANSWERED on the 19th of April AWG meeting

3b) Background processes @ 3 TeV - qqll$\nu\nu$

- **Two processes have ben emphasized:**
  - **qqqqll**
  - **qqllvv at 3 TeV**

- This background is not produced by the Collaboration (MonteCarloSamplesForTheHiggsPaper)
- We've produced 10.000 events in Whizard v1.95. to see how it behaves
- $\sigma(e^+e^- \rightarrow q\bar{q}ll\nu\bar{\nu}) \sim$ 3.4 fb (in 5 ab^{-1}, ~17000 events).
- Signal: $\sigma(e^+e^- \rightarrow H\nu\nu, H \rightarrow ZZ^* \rightarrow q\bar{q}ll, l = e^\pm, \mu^\pm)$, σ = 1.13 fb, in 5 ab^{-1}, 5650 events.

- **Just by requiring Higgs mass window** for invariant mass of the P(qq)+P(ll) four-vector sum (90 GeV < m_H < 160 GeV), **background drops 85 times** (i.e. 17000→200 events)

- Cuts on available MVA variables at the generated level (established in the training without qqllvv background): $m_Z, m_{Z^*}, m_{l^+l^-}, m_{q\bar{q}}, m_H, E_{vis}, E_{vis} - E_H, p_T^{miss}$, θ_H, N_{PFO}, **reduce background completely.**

- Even with the single cut on a Higgs mass, with N_S = 2281, N_B = 2979(+200), statistical uncertainty will be 3.24% instead of previous 3.18%.

19th of April AWG meeting, comments/questions raised:

1) Implement lepton dressing @ 350 GeV and 3 TeV for background processes as well (P. Roloff)
2) *Re-weight reconstructed electron IP distributions according to the ones for muons (P. Roloff)*

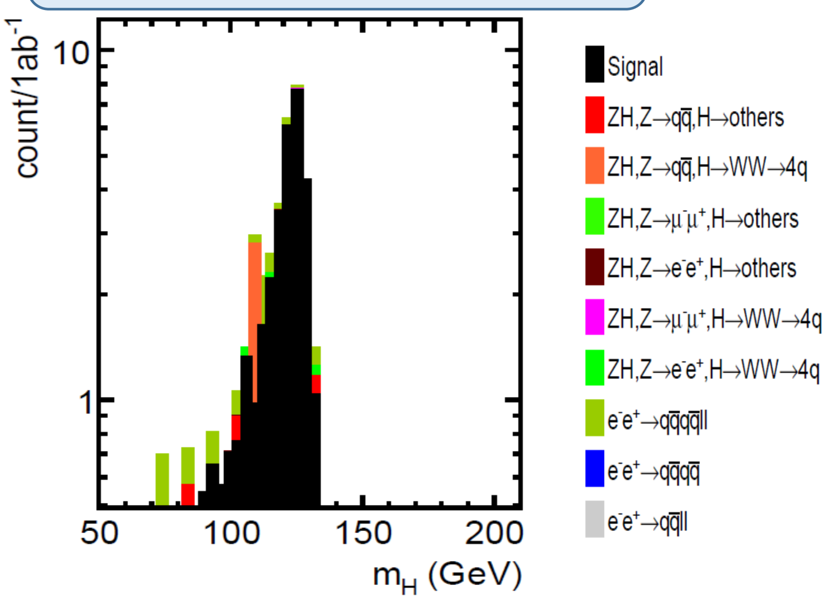
Discussion with opposite stand points of physics conveners. Authors consider such a reweight inadequate as this would not have physical justification and may even be wrong, due to the fact that electrons from Z decays can undergo multiple scattering much more than muons and much more than i.e. spectator electrons from ZZ-fusion Higgs decay to ZZ$\rightarrow$qqee.

This will turned out true in the reconstructed electron IP distributions, where IP enlargement is seen with one and not seen with the other type of processes, using the same detector model (see the last slide).

1) ANSWERED on the 17th of May AWG meeting

@ 350 GeV

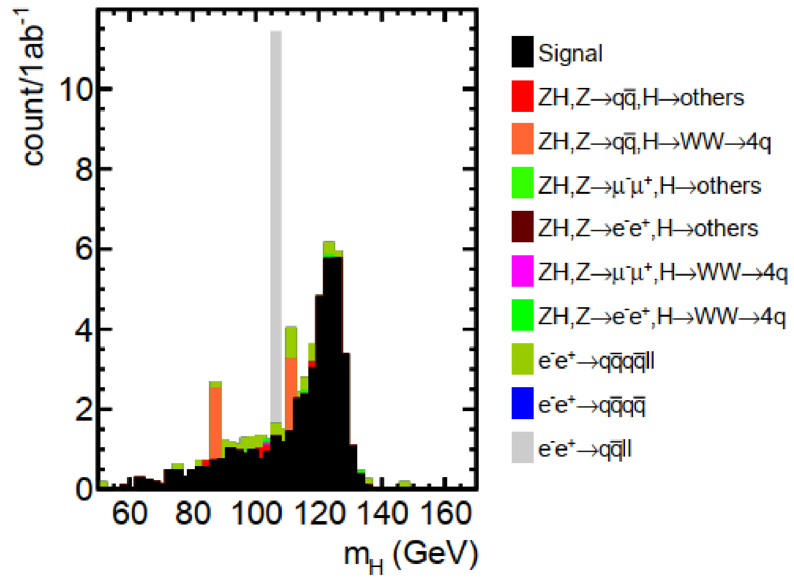
only signal is dressed



δ ≈ 18.1%

Process	ε_{presel}	ε_{BDT}	N_{BDT}
Signal @ 350 GeV	77%	19%	35
Background process			
$e^-e^+ \to q\bar{q}q\bar{q}l^+l^-$	21%	3‰	3
$e^-e^+ \to HZ; Z \to q\bar{q}, H \to WW \to 4q$	0.42%	4.5%	2

signal and background dressed



δ = 18.3%

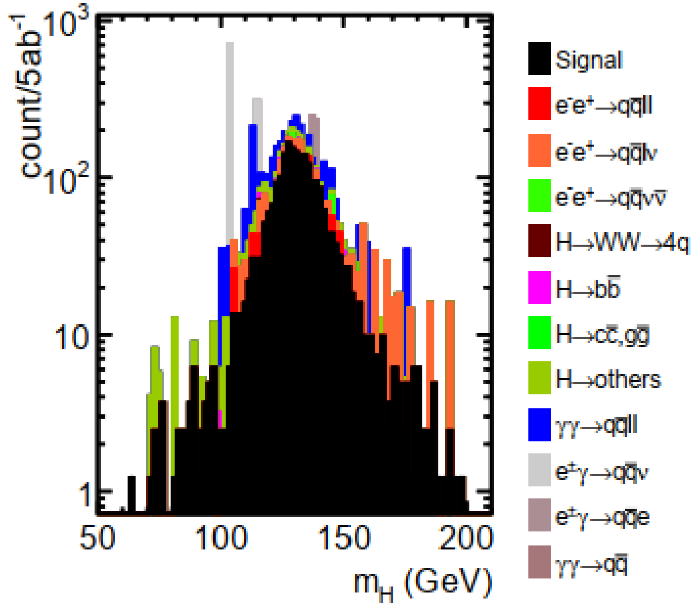
Process	ε_{presel}	ε_{BDT}	N_{BDT}
Signal @ 350 GeV	77%	23%	43
Background process			
$e^-e^+ \to q\bar{q}l^+l^-$	11.4%	0.05‰	10
$e^-e^+ \to q\bar{q}q\bar{q}l^+l^-$	21%	5‰	5
$e^-e^+ \to HZ; Z \to q\bar{q}, H \to WW \to 4q$	0.42%	9%	4

NEW

1) ANSWERED on the 17th of May AWG meeting

@ 3 TeV

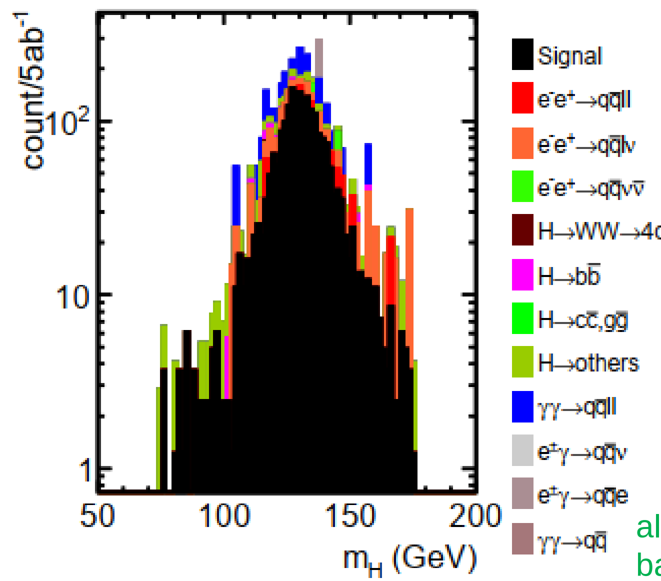
only signal is dressed



δ = 3.2%

Process	ε_{presel}	ε_{BDT}	N_{BDT}
Signal @ 3 TeV	67%	60%	2281
Background process			
$e^{\pm}\gamma \to q\bar{q}\nu$	1.4‰	5‰	930
$\gamma\gamma \to q\bar{q}l^+l^-$	11‰	0.5‰	820
$e^-e^+ \to q\bar{q}l\nu$	3‰	6‰	482
$e^-e^+ \to H\nu\bar{\nu}; H \to others$	45‰	2%	345
$e^{\pm}\gamma \to q\bar{q}e$	8.8‰	0.4‰	232
processes with N_{BDT}<100	5.2‰	1.3‰	170

signal and background dressed



almost 40% background less

δ = 2.9%

NEW

Process	ε_{presel}	ε_{BDT}	N_{BDT}
Signal @ 3 TeV	67%	59%	2232
Background process			
$\gamma\gamma \to q\bar{q}l^+l^-$	11‰	0.4‰	672
$e^-e^+ \to q\bar{q}l\nu$	3‰	6‰	509
$e^-e^+ \to H\nu\bar{\nu}; H \to others$	45‰	1.6%	328
$e^-e^+ \to q\bar{q}l^+l^-$	7.5‰	1‰	126
$e^{\pm}\gamma \to q\bar{q}e$	8.8‰	0.2‰	116
processes with N_{BDT}<100	0.7‰	1.7%	98

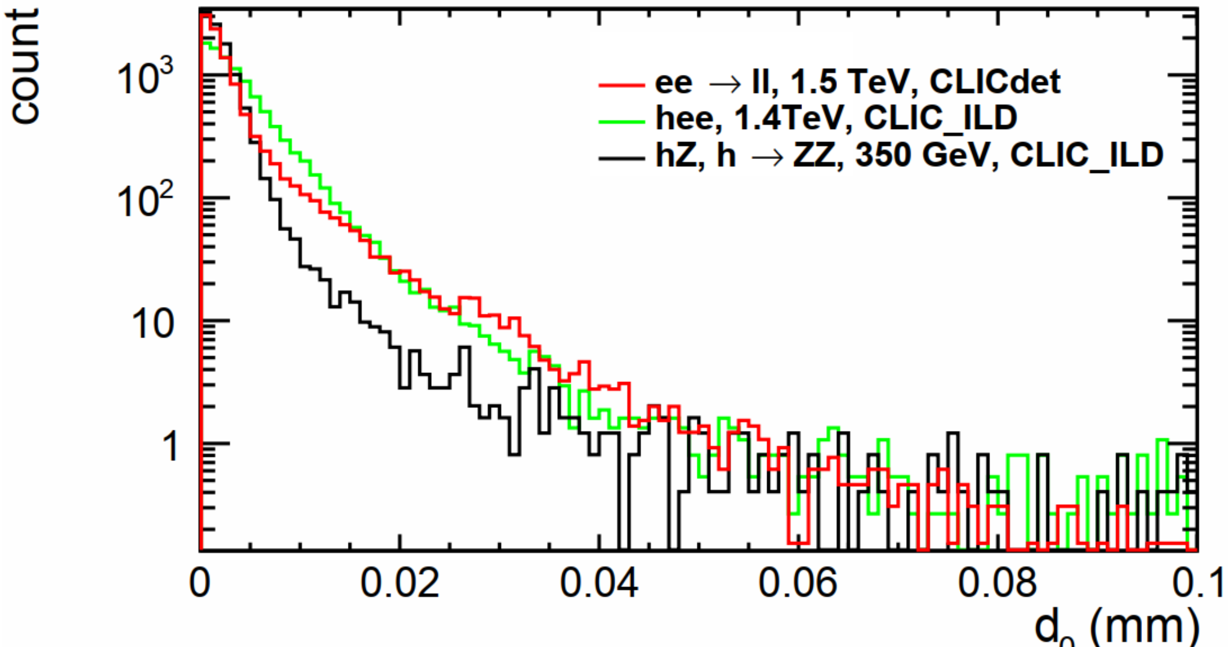
17^{th} of May AWG meeting, comments/questions raised:

1) Try to make comparison among e^- transverse IP for the processes produced in CLICdet and CLIC_ILD geometry (F.Zarnecki)

ANSWERED to F. Zarnecki on the 7th of June 2021

e⁻ IP in CLICdet/CLIC_ILD geometry

- Comparison among e⁻ transverse IP (d_0) is made for processes: hee @ 1.4 TeV and hZ, h $\rightarrow ZZ^*$ @ 350 GeV, both produced in CLIC_ILD geometry and $ee \rightarrow ll$ @ 1.5 TeV, produced in new (CLICdet) geometry



All distributions are normalized to the same integral - 10000 events

- It can be seen both detectors (red and green) behave similarly with 'the same type of electrons' (i.e produced directly from 2nd out of 6 vertices)
- Loss of events with electrons in the case of $H \rightarrow ZZ^* \rightarrow q\bar{q}ll$ at 350 GeV given in black, might be a consequence of kinematics making the final state electrons more prone to multiple scattering
- There was no similar process to $H \rightarrow ZZ^* \rightarrow q\bar{q}ll$ available with the CLICdet

Continuation July/August

Comments/questions to the analysis raised after the 19th of July by F.Zarnecki:

Q: line 265: "P(b) and P(c) are b-tag and c-tag probabilities of a randomly chosen jet" - this seems a very strange way to do it. If you do expect flavour tagging to be important and do not want to include 4x2 tag values, you can use their sums. It would look much better than selecting at random...

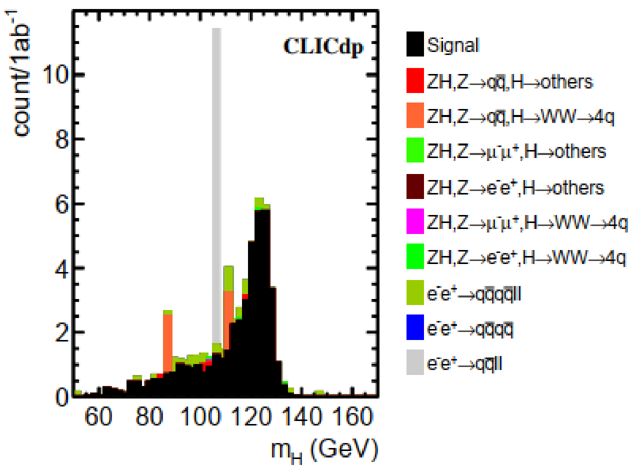
line 299: "where 1 and 2 denote randomly chosen jets," - you do consider both jets here. It makes much more sense to order jets in pt or in P(b) than to take it randomly. You can only improve your selection this way...

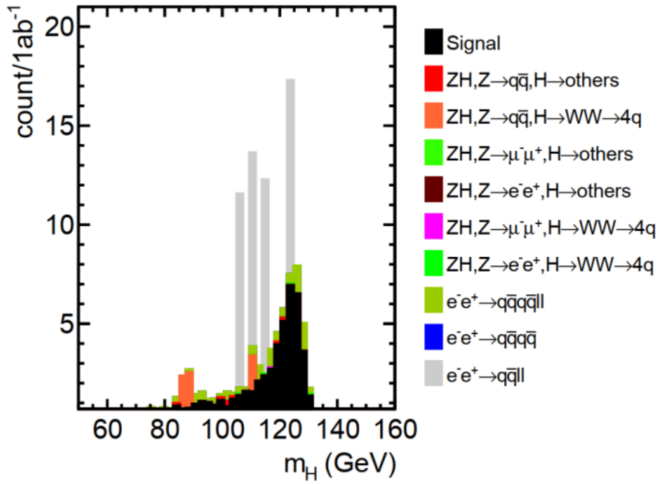
A: a) We have only 2 jets (from the Higgs decays to ZZ*) – not 4.

b) Flavor tagging probabilities are among least sensitive (worst MVA ranking)

c) We've introduced pt ordering

MVA selection @ 350 GeV without and with pt ordering of $P_{b,c}$





δ = 18.3%

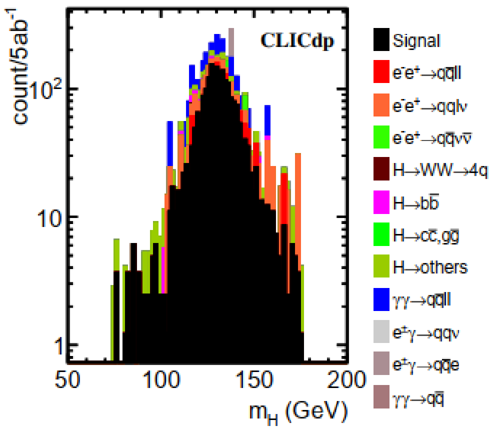
Process	ε_{BDT}	N_{BDT}
Signal @ 350 GeV	23%	43
Background process		
$e^-e^+ \to q\bar{q}l^+l^-$	0.05‰	10
$e^-e^+ \to q\bar{q}q\bar{q}l^+l^-$	5‰	5
$e^-e^+ \to HZ; Z \to q\bar{q}, H \to WW \to 4q$	9%	4

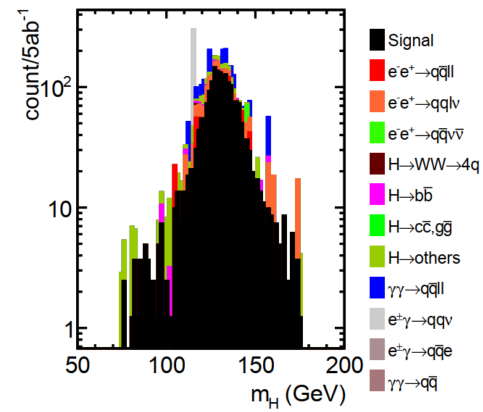
δ = 19.7%

Process	ε_{BDT}	N_{BDT}
Signal @ 350 GeV	29%	53
Background process		
$e^-e^+ \to q\bar{q}l^+l^-$	0.2‰	39
$e^-e^+ \to q\bar{q}q\bar{q}l^+l^-$	1.2%	11
$e^-e^+ \to HZ; Z \to q\bar{q}, H \to WW \to 4q$	11.4%	5
$e^-e^+ \to HZ; Z \to q\bar{q}, H \to WW \to others$	3.8%	1

There is no meaningful interpretation/impact, since i.e. instead of 1 we have 4 (not-normalized) qqll events with a scaling factor ~10…

MVA selection @ 3 TeV without and with pt ordering of $P_{b,c}$





δ = 3.1%

Process	ε_{BDT}	N_{BDT}
Signal @ 3 TeV	59%	2232
Background process		
$\gamma\gamma \to q\bar{q}l^+l^-$	0.4‰	672
$e^-e^+ \to qql\nu$	6‰	509
$e^-e^+ \to H\nu\bar{\nu}; H \to others$	1.6%	328
$e^-e^+ \to q\bar{q}l^+l^-$	1‰	126
$e^{\pm}\gamma \to q\bar{q}e$	0.2‰	116
processes with N_{BDT}<100	1.7%	98

δ = 2.9%

Process	ε_{BDT}	N_{BDT}
Signal @ 3 TeV	53%	2020
Background process		
$\gamma\gamma \to q\bar{q}l^+l^-$	0.3‰	438
$e^-e^+ \to q\bar{q}l\nu$	4‰	322
$e^-e^+ \to H\nu\bar{\nu}; H \to others$	1.3%	259
$e^{\pm}\gamma \to q\bar{q}\nu$	1.3‰	252
processes with N_{BDT}<100	1.1‰	140

Slight reduction (~10%) of background, but also less signal efficiency. Uncertainty is ~3%
We believe this behavior (also at 350 GeV) is due to the fact that flavor tagging variables are not highly ranked in any case.

Comments/questions to the analysis raised after the 19th of July (P.Roloff):

Q/Comment: The range of the X-axis (d_0) needs to be extended as it did not show the region where the problem in the analysis was visible.
Muons need to be added for comparison.
It would be good to also show z_0.

There are samples available much more suitable for this test like:

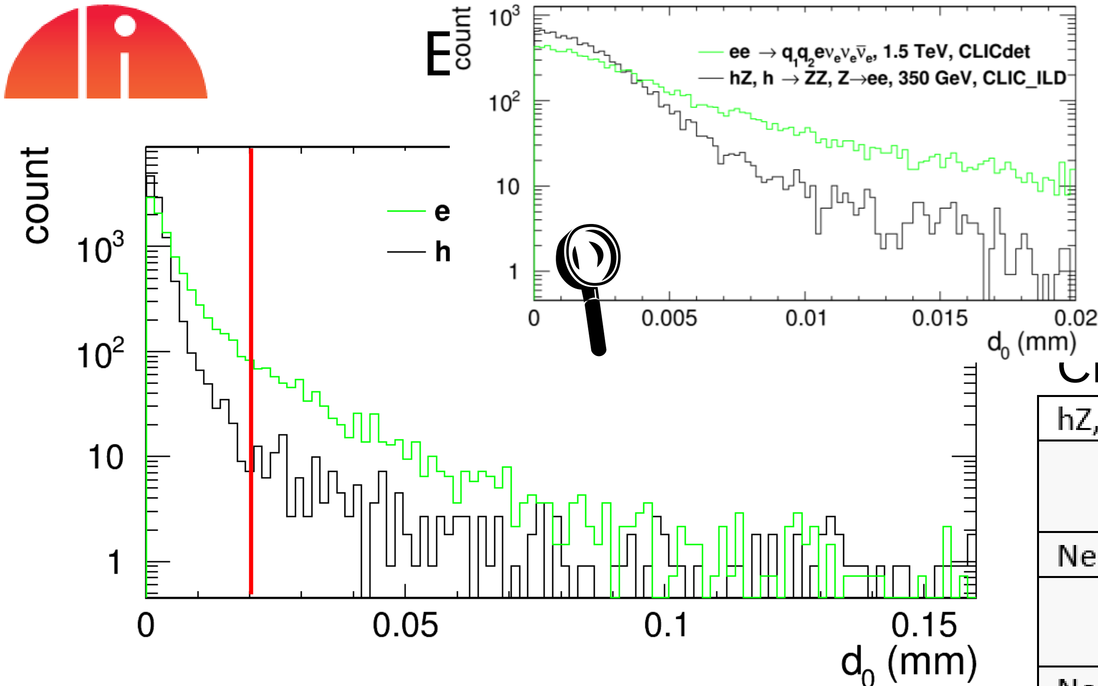
ee->qqlvvv

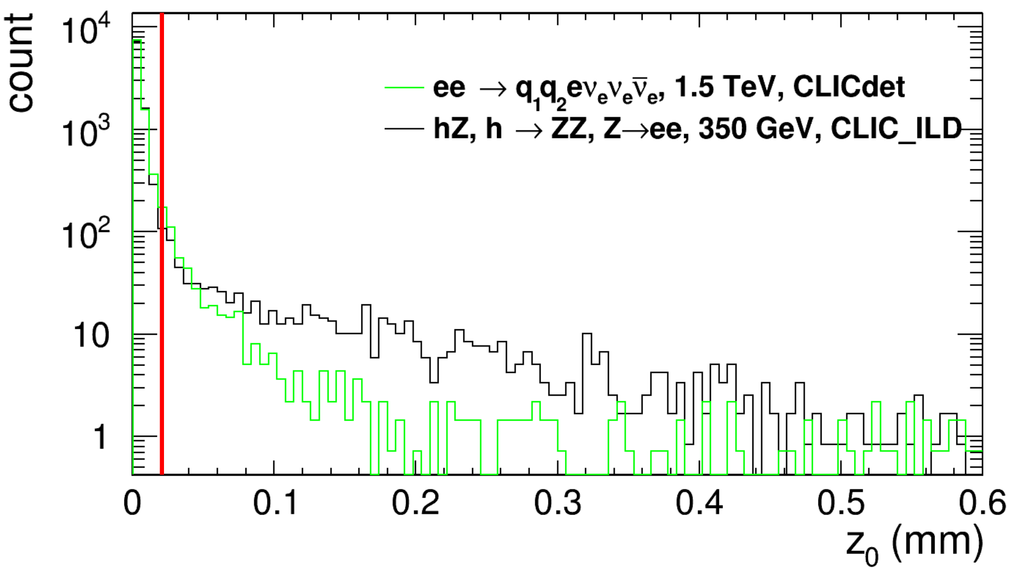
A: a) x-scale max of ~0.1 mm for d_0 is meaningful as we cut at 0.02mm and nothing really happens above 0.1mm (see the next slide).

b) we use now ee->qqlvvv for comparison

c) add z_0 and muons

E ers 10k particles

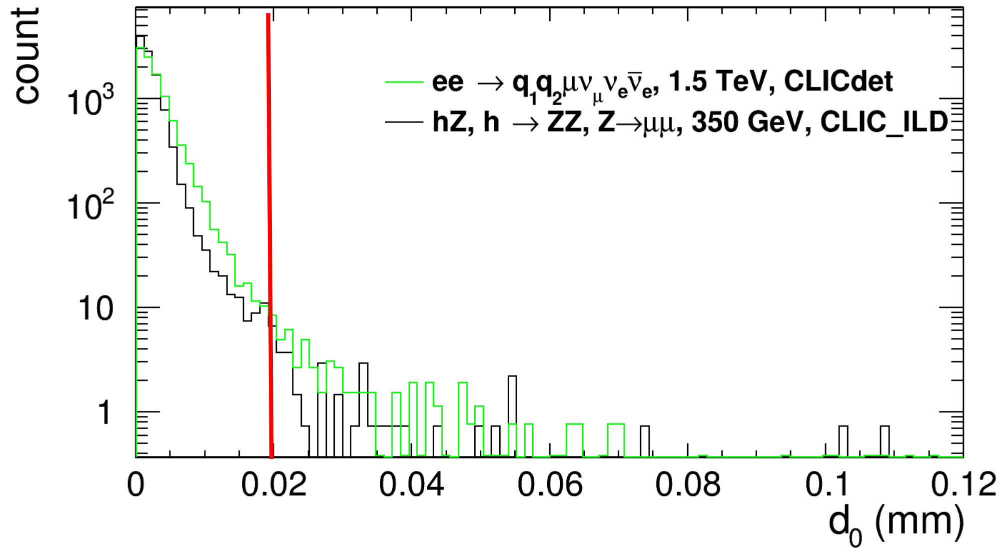




CLIC_ILD	CLICdet
hZ, h→ZZ, Z→ee, 350 GeV	ee→$q\bar{q}e\nu\nu\nu$, 1.5 TeV
$d_0 < 0.02$	
Ne = 9150, 8.5%	9152, 8.5%
$z_0 < 0.02$	
Ne = 9080, 9.2%	9390, 6.1%
$z_0 < 0.02$ && $d_0 < 0.02$	
Ne = 8910, 11%	8751, 12.5%

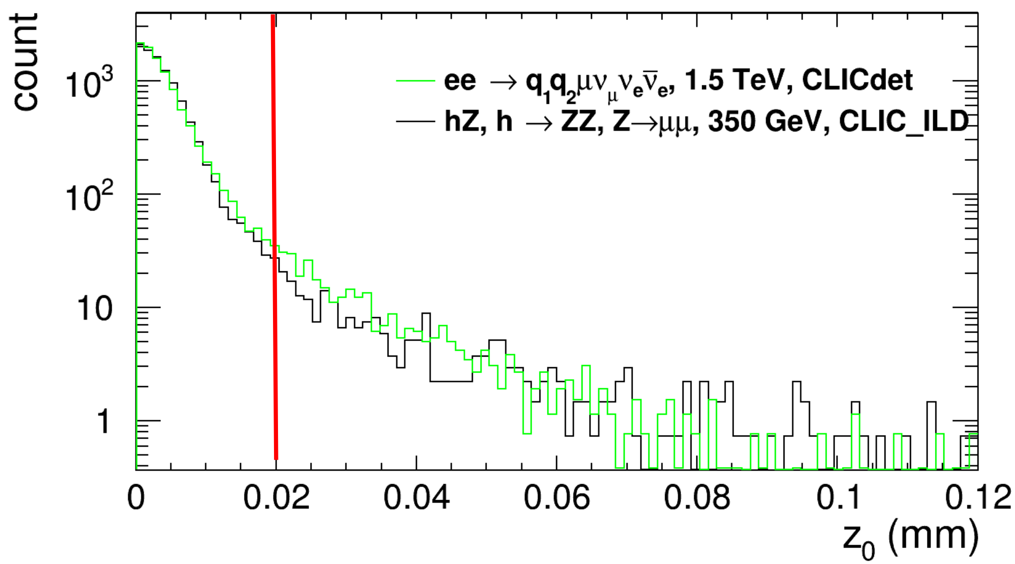
NO SIGNIFICANT DIFFERENCE IN DETECTOR PERFORMANCE

Muon impact parameters 10k particles



CLIC_ILD CLICdet

hZ, h→ZZ, Z→ μμ, 350 GeV	ee→q$\bar{q}$μνν, 1.5 TeV
$d_0 < 0.02$	
Nmu = 9970, 0.3%	9939, 0.6%
$z_0 < 0.02$	
Nmu = 9720, 2.8%	9642, 3.6%
$z_0 < 0.02$ && $d_0 < 0.02$	
Nmu = 9720, 2.8%	9613, 3.9%



NO SIGNIFICANT DIFFERENCE IN DETECTOR PERFORMANCE
Muon impact parameters are ~2 times smaller

As the compared processes behave similarly, the sentence from line 68 about models' performances vs. IP reconstruction will be removed (authors' suggestion)