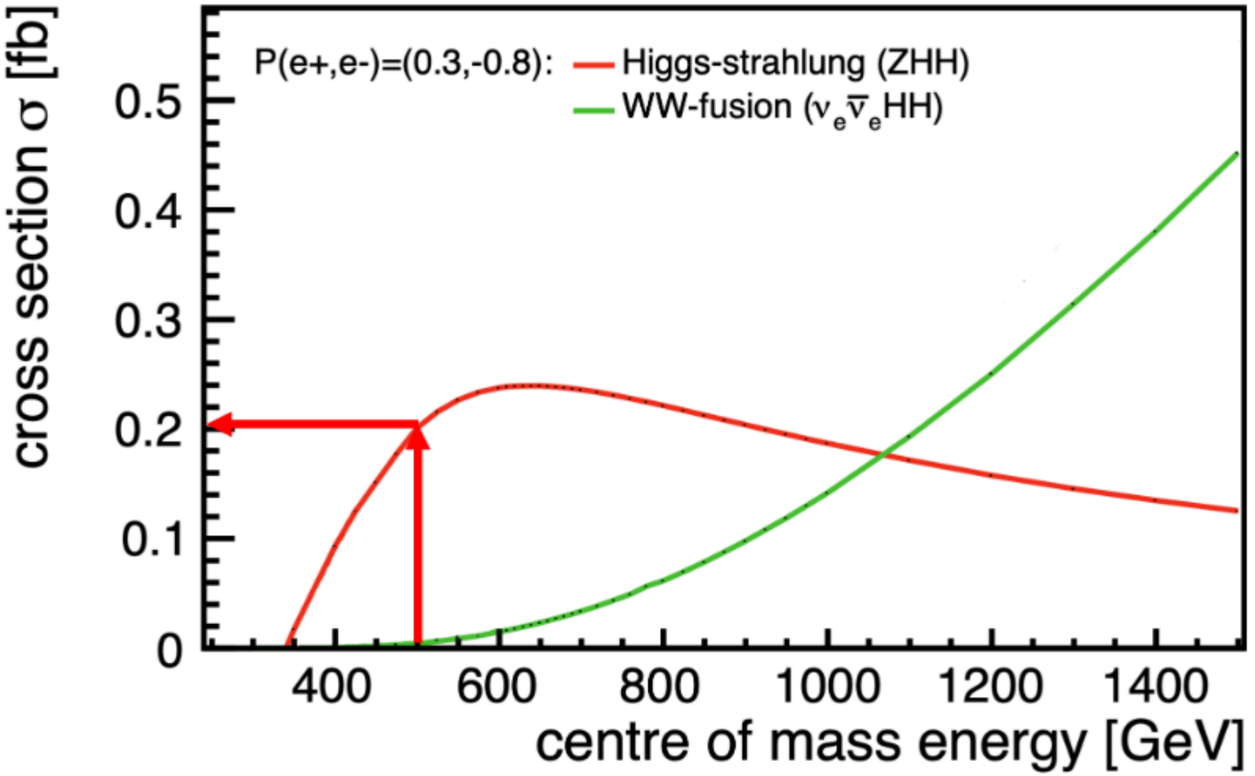
Update on new Higgs self-coupling study

And proposal for MC@550 GeV

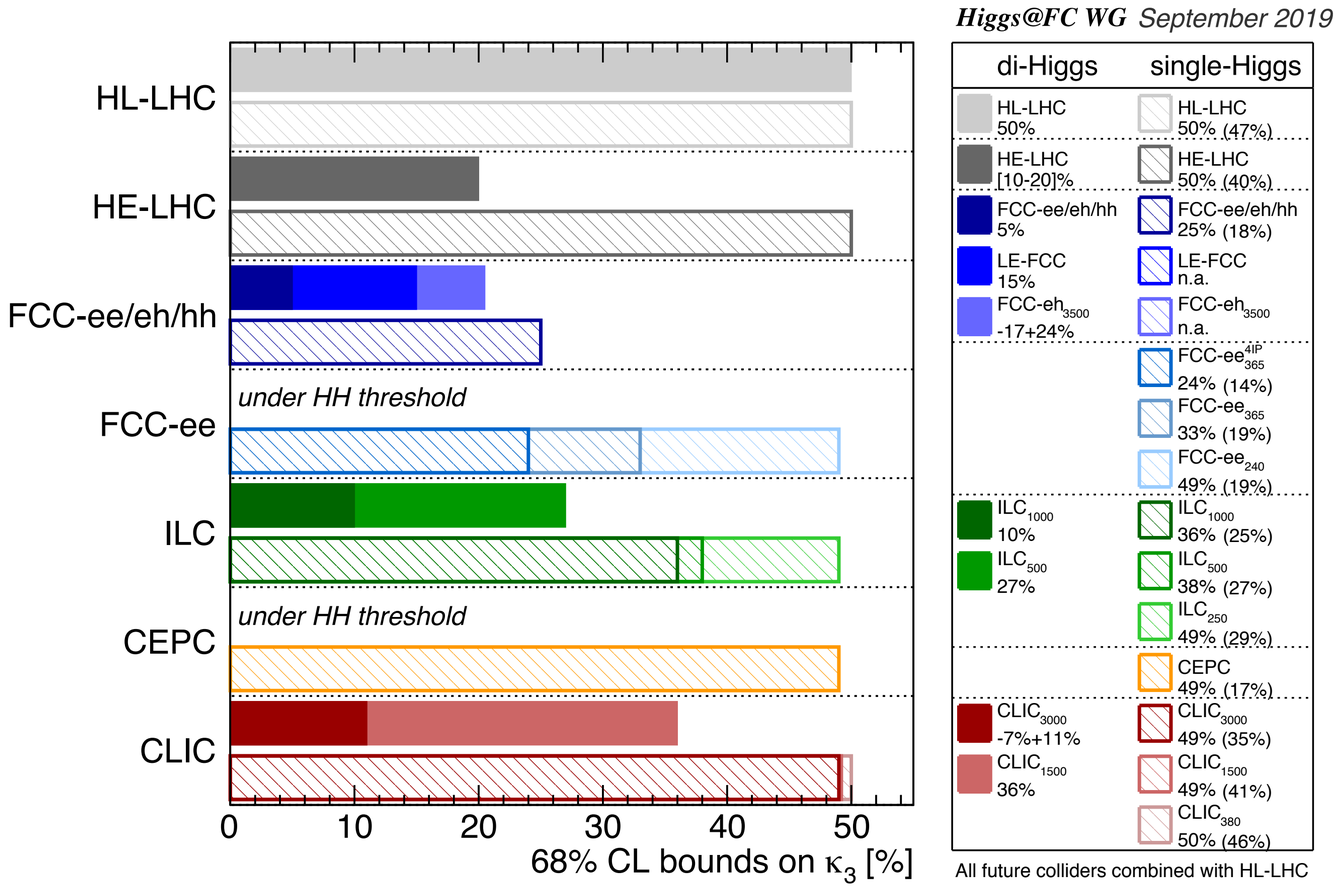


Jenny List (DESY)

ILD Software & Analysis Meeting, Jun 5 2024

Motivation I

Higgs self-coupling as key part of physics case for e+e- collisions at >= 500 GeV



Basic argument of circular collider community:

- don't need ECM > 350 GeV since FCChh will do much better on Higgs self-coupling than high-E e+e-

As we all know, this is comparing apples vs bananas:

- **fast sim of a detector one doesn't know how to build**
  vs
  ILD full sim from 10 years ago!

Motivation II

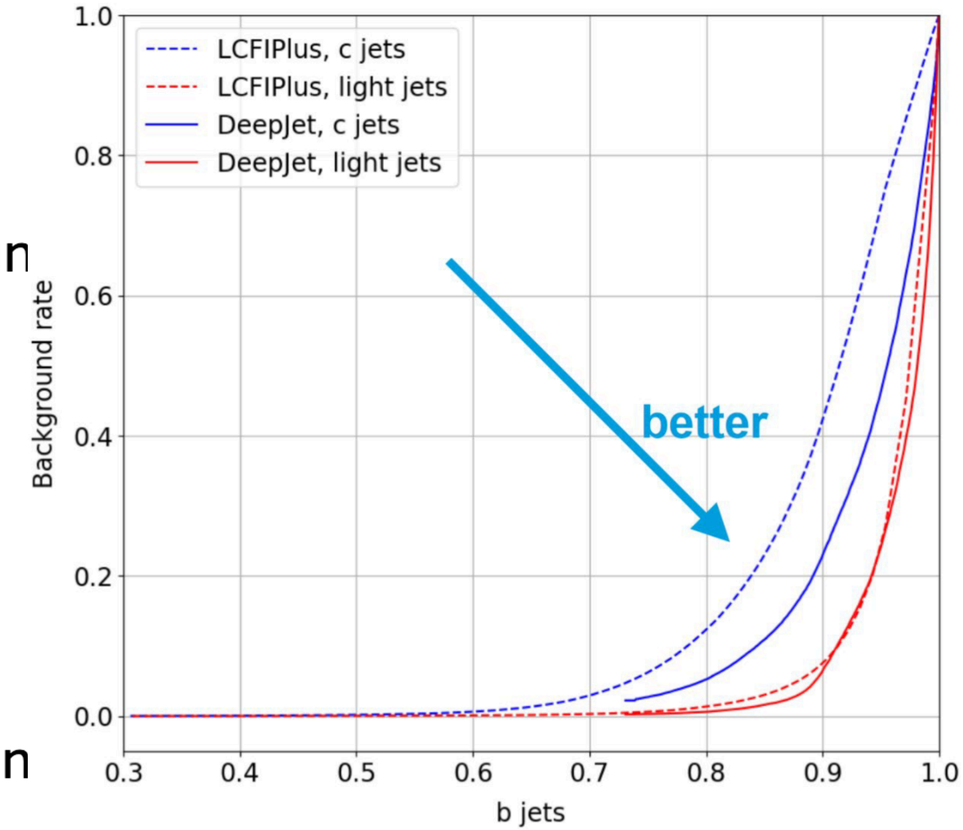
Some estimated (relative) improvements since PhD thesis of Claude Dürig

- jet pairing and jet misclustering: “perfect“ jet clustering → 40% improvement
  improve di-jet mass resolution
- removal of $\gamma\gamma$ overlay: 15% improvement expected
  important to tackle initial state radiation (ISR)
- flavor tagging: 11% improvement expected from 5% eff. increase with newer LCFIPlus
  important as $H \to b\bar{b}$ is the dominant Higgs decay channel
- adding $Z \to \tau\tau$ channel: 8% improvement expected
  include a yet unaccounted decay channel
- tagging of isolated leptons
  improves reconstruction of Z bosons
- separation of ZHH diagrams with/without the self-coupling
  would directly improve the sensitivity on λ (lower sensitivity factor)

Motivation II

Some estimated (relative) improvements since PhD thesis of Claude Dürig

- jet pairing and jet misclustering: “perfect“ jet clustering → 40% improvement
  improve di-jet mass resolution
- removal of $\gamma\gamma$ overlay: 15% improvement expected
  important to tackle initial state radiation (ISR)
- flavor tagging: 11% improvement expected from 5% eff. in
  important as $H \to b\bar{b}$ is the dominant Higgs decay channel
- adding $Z \to \tau\tau$ channel: 8% improvement expected
  include a yet unaccounted decay channel
- tagging of isolated leptons
  improves reconstruction of Z bosons
- separation of ZHH diagrams with/without the self-couplin
  would directly improve the sensitivity on λ (lower sensitivity factor)



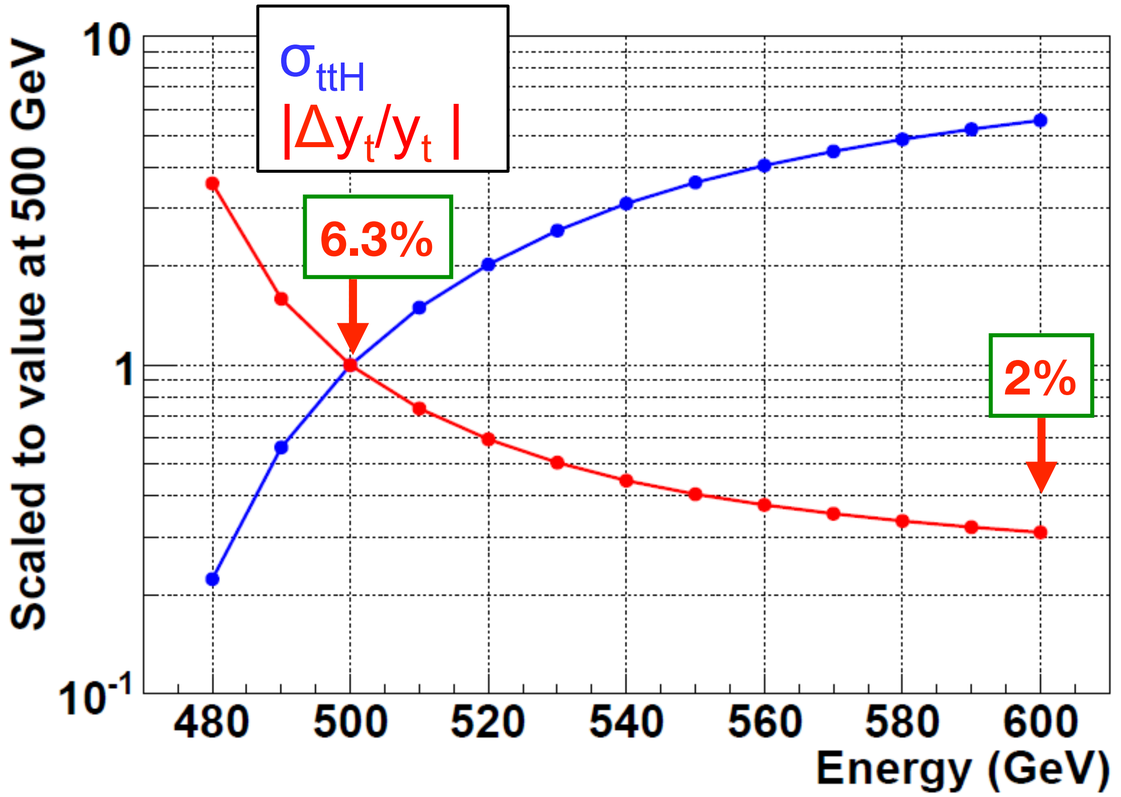
Flavor tagging performance of LCFIPlus vs. DeepJet at ILD full simulation.
M. Meyer [2023]

Motivation III

C3 decided for 550 GeV & IDT will cost ILC 250 / 350 / 550 (!)

- known since long:
  ttH strongly prefers ECM larger than 500 GeV
- impact on **ZHH** less clear:
  - **cross-section rises**
  - but **relative sensitivity to λ drops**
    (i.e. cross-section growth from diagrams not depending λ)
  - higher boost: **facilitates b-tagging, jet clustering ?**
  - is there an optimum?

=> need to try out!



The Proposal - from fall 2021

Overview

- **event generation:**
  - start with only “ZHH” and “ZZH” (as usual each Z decay mode, separated into qq, ll, ee, vv)
  - 3 ECMs: 500 GeV, 550 GeV, 600 GeV
  - for aim ~ 2x the statistics Claude had (per ECM) => ~6 M events per ECM, i.e. 18M total
- **sim/rec:**
  - start with 500 GeV (for comparison to Claude approx. compatibility with IDR samples
  - eventually also 550 or 600 GeV, t.b.d. after generator-level (or SGV-level?) comparisons

The Proposal - from fall 2021

Overview

- **event generation:**
  - start with only “ZHH” and “ZZH” (as usual each Z decay mode, separated into qq, ll, ee, vv)
  - 3 ECMs: 500 GeV, 550 GeV, 600 GeV
  - for aim ~ 2x the statistics Claude had (per ECM) => ~6 M events per ECM, i.e. 18M total
- **sim/rec:**
  - start with 500 GeV (for comparison to Claude approx. compatibility with IDR samples
  - eventually also 550 or 600 GeV, t.b.d. after generator-level (or SGV-level?) comparisons

The Proposal - from fall 2021

Overview

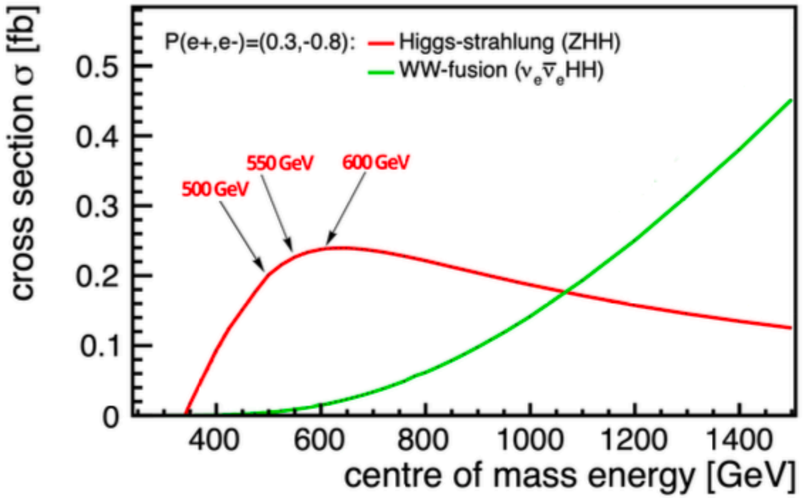
- **event generation:**
  - start with only “ZHH” and “ZZH” (as usual each Z decay mode, separated into qq, ll, ee, vv)
  - 3 ECMs: 500 GeV, 550 GeV, 600 GeV
  - for aim ~ 2x the statistics Claude had (per ECM) => ~6 M events per ECM, i.e. 18M total

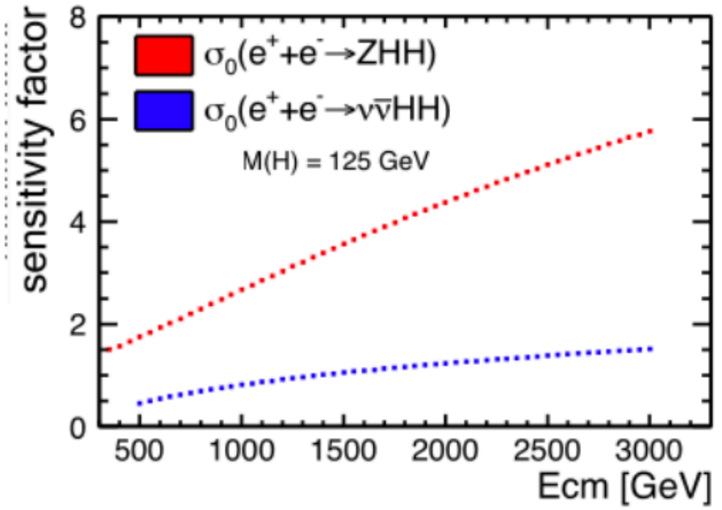
- **sim/rec:**
  - start with 500 GeV (for comparison to Claude approx. compatibility with IDR samples
  - eventually also 550 or 600 GeV, t.b.d. after generator-level (or SGV-level?) comparisons

ZHH and ZZH @ 500 / 550 / 600 GeV

produced 2022/23 - compared by Julie Torndal, cf her presentation at EPS-HEP2023
https://indico.desy.de/event/34916/contributions/147294/

- More ZHH events **but** larger σ_{ZHH} contributions from diagrams NOT containing the Higgs self-coupling





Advantages of going to higher energies:

- More boosted jets

→ Less misclustering, better jet-pairing?

→ Improved b-tagging efficiencies?

→ Better kinematic separation of signal and background?

Disadvantages of going to higher energies:

- Sensitivity factor increases with the E_{CM}

→ Less sensitivity to Higgs self-coupling?

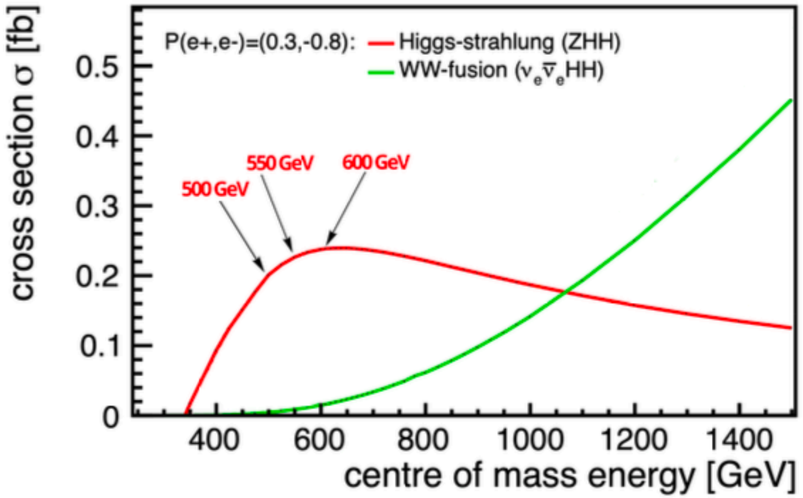
$$\frac{\Delta\lambda}{\lambda} = c \cdot \frac{\Delta\sigma}{\sigma}$$

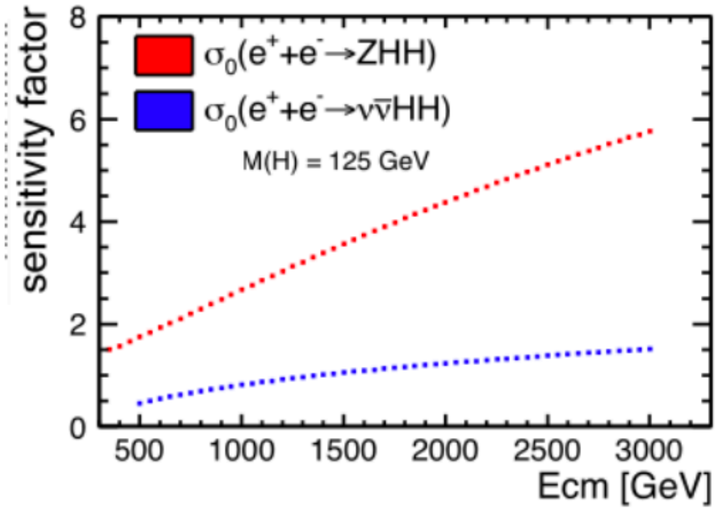
c: sensitivity factor

ZHH and ZZH @ 500 / 550 / 600 GeV

produced 2022/23 - compared by Julie Torndal, cf her presentation at EPS-HEP2023
https://indico.desy.de/event/34916/contributions/147294/

- More ZHH events **but** larger σ_{ZHH} contributions from diagrams NOT containing the Higgs self-coupling





Advantages of going to higher energies:

- More boosted jets

→ Less misclustering, better jet-pairing?

→ Improved b-tagging efficiencies?

→ Better kinematic separation of signal and background?

Disadvantages of going to higher energies:

- Sensitivity factor increases with the E_{CM}

→ Less sensitivity to Higgs self-coupling?

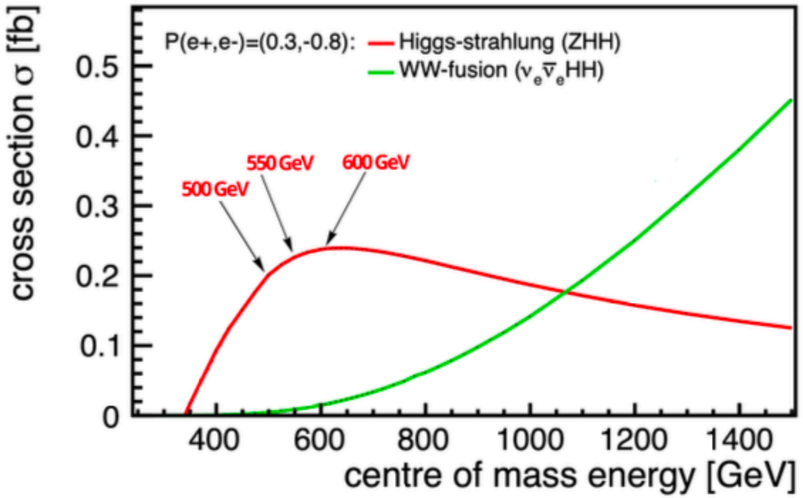
$$\frac{\Delta\lambda}{\lambda} = c \cdot \frac{\Delta\sigma}{\sigma}$$

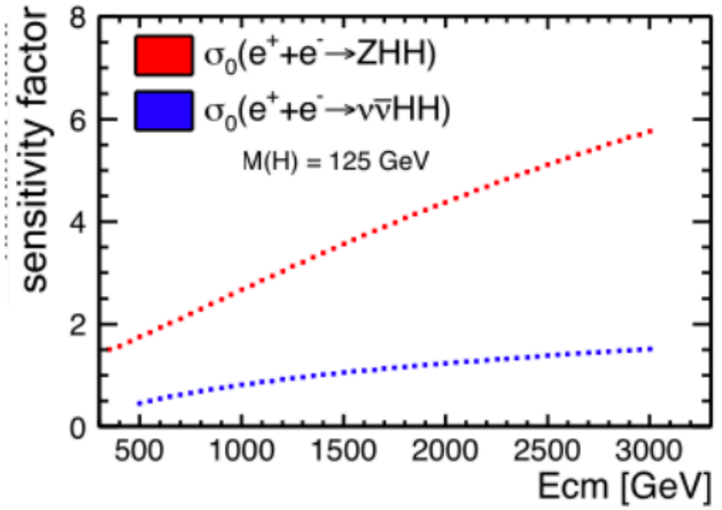
c: sensitivity factor

ZHH and ZZH @ 500 / 550 / 600 GeV

produced 2022/23 - compared by Julie Torndal, cf her presentation at EPS-HEP2023
https://indico.desy.de/event/34916/contributions/147294/

- More ZHH events **but** larger σ_{ZHH} contributions from diagrams NOT containing the Higgs self-coupling





c (sigma-tot-only) changes from 1.9 @500 GeV to 1.95 @550 GeV => minor effect

Advantages of going to higher energies:

- More boosted jets

→ Less misclustering, better jet-pairing?

→ Improved b-tagging efficiencies?

→ Better kinematic separation of signal and background?

Disadvantages of going to higher energies:

- Sensitivity factor increases with the E_{CM}

→ Less sensitivity to Higgs self-coupling?

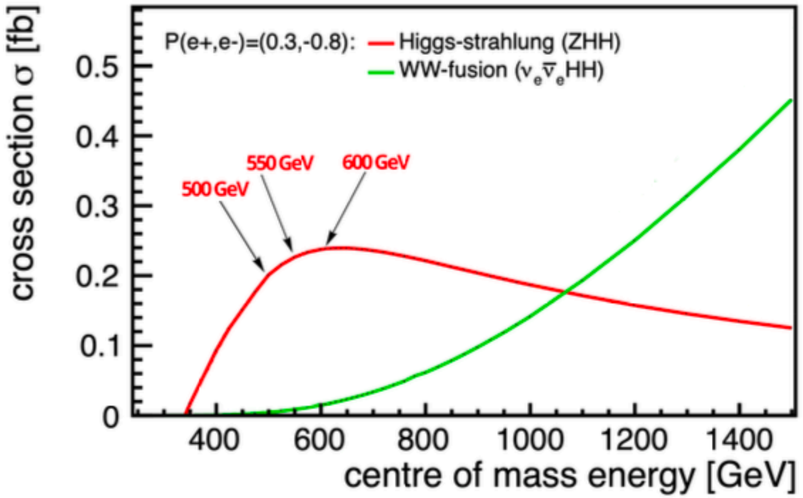
$$\frac{\Delta\lambda}{\lambda} = c \cdot \frac{\Delta\sigma}{\sigma}$$

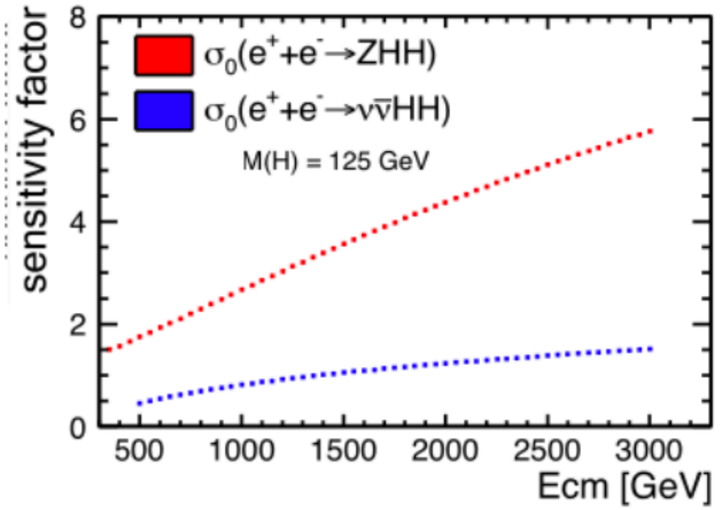
c: sensitivity factor

ZHH and ZZH @ 500 / 550 / 600 GeV

produced 2022/23 - compared by Julie Torndal, cf her presentation at EPS-HEP2023
https://indico.desy.de/event/34916/contributions/147294/

- More ZHH events **but** larger σ_{ZHH} contributions from diagrams NOT containing the Higgs self-coupling





c (sigma-tot-only) changes from 1.9 @500 GeV to 1.95 @550 GeV => minor effect

✔

Advantages of going to higher energies:

- More boosted jets

→ Less misclustering, better jet-pairing?

→ Improved b-tag[illegible] efficiencies?

→ Better kinematic separation of signal and background?

✔

Disadvantages of going to higher energies:

- Sensitivity factor increases with the E_{CM}

→ Less sensitivity to Higgs self-coupling?

$$\frac{\Delta\lambda}{\lambda} = c \cdot \frac{\Delta\sigma}{\sigma}$$

c: sensitivity factor

DESY.

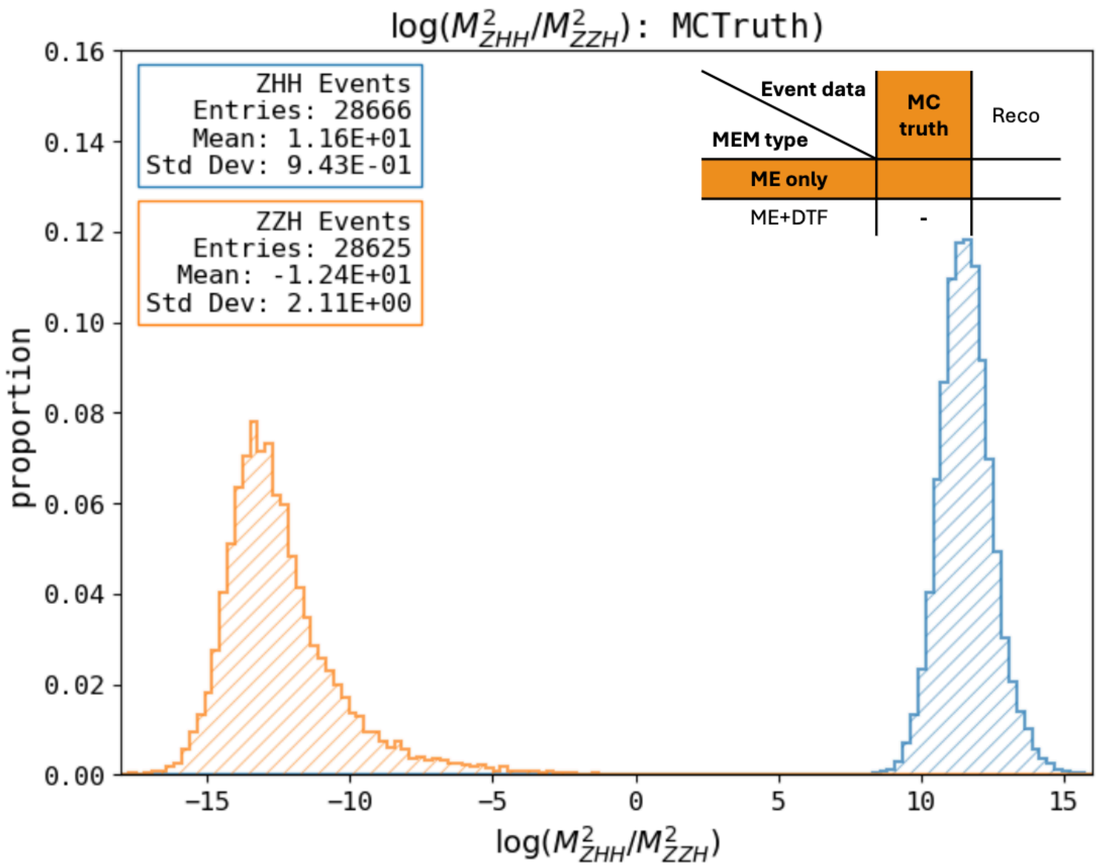
More analysis improvements

from 2023/24 - by Bryan Bliewert, cf his recent presentation at ECFA Hself Focus Meeting

https://indico.cern.ch/event/1413943/

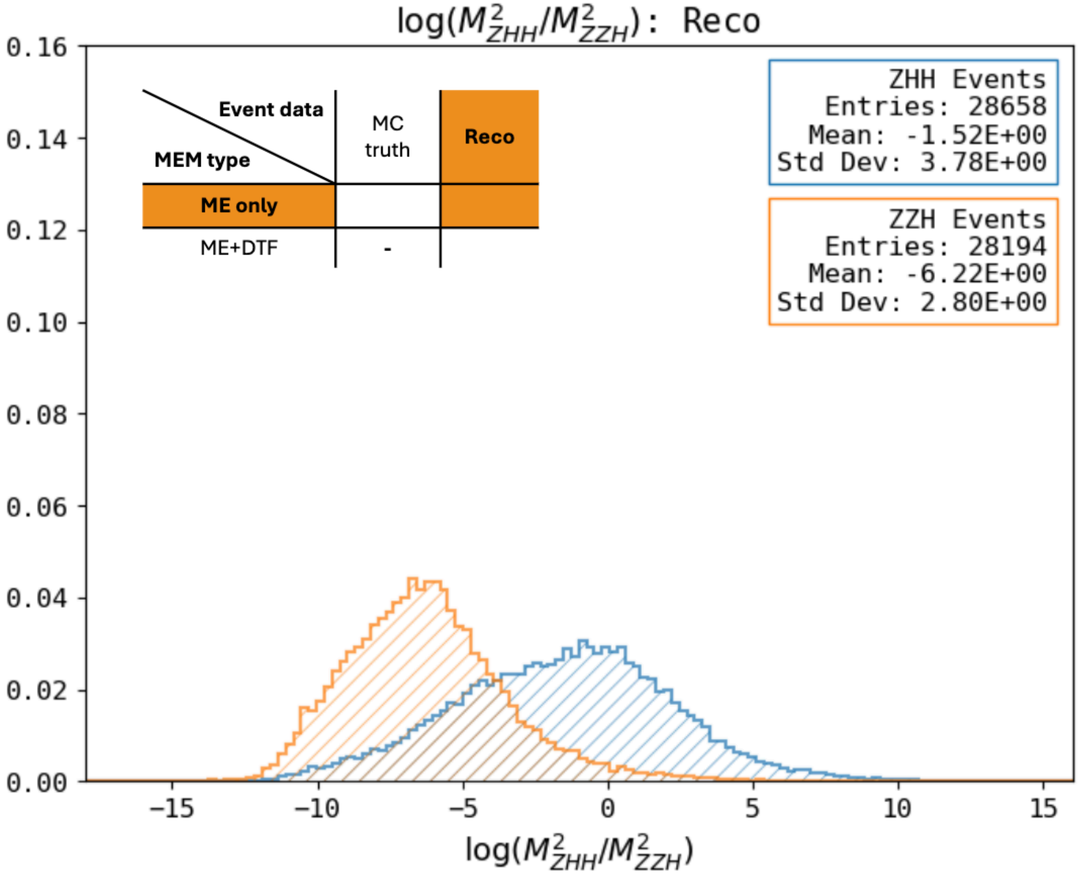
generator level check

- excellent separation



naive MEM

- significant separation power lost

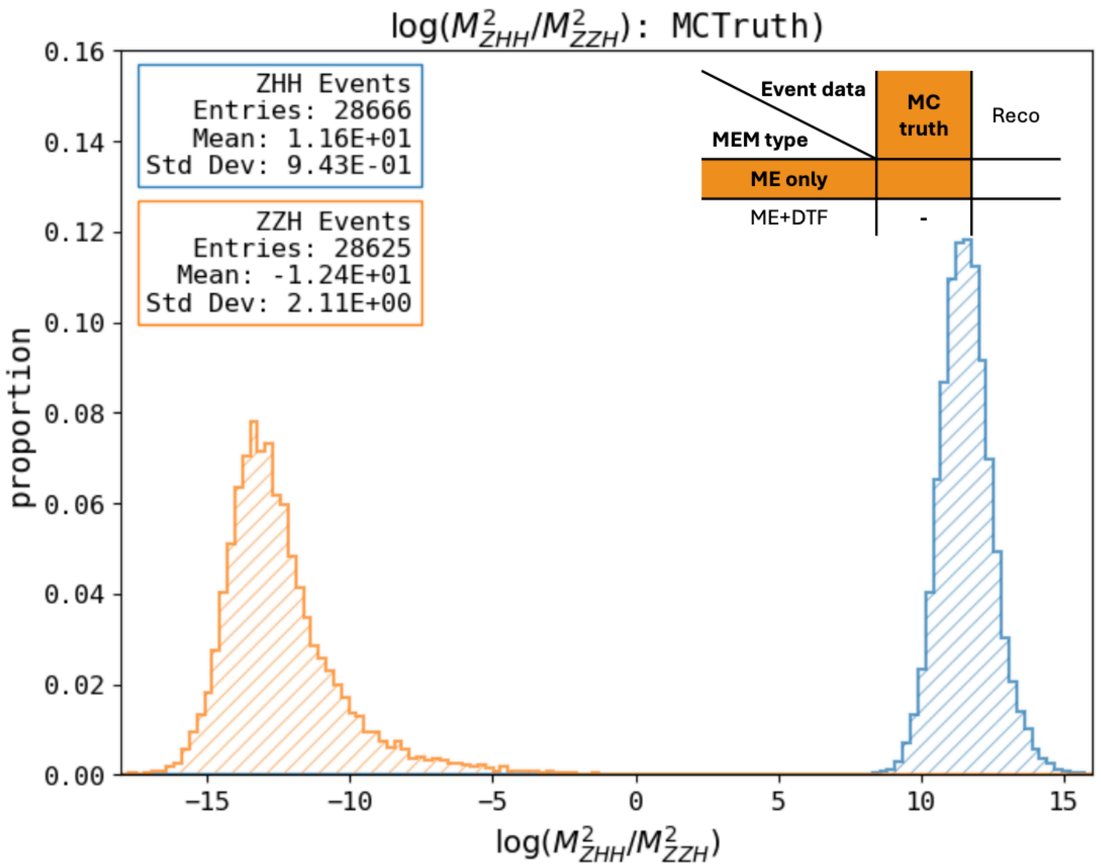


More analysis improvements

from 2023/24 - by Bryan Bliewert, cf his recent presentation at ECFA Hself Focus Meeting
https://indico.cern.ch/event/1413943/

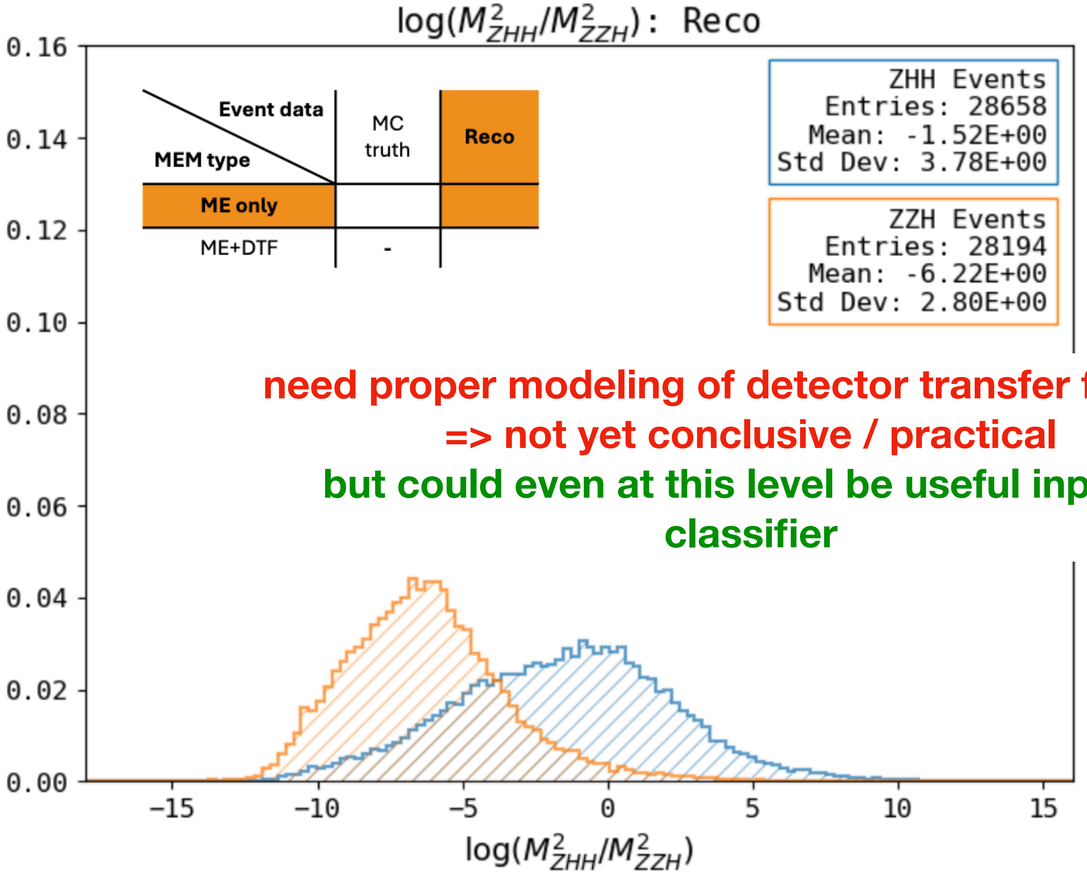
generator level check

- excellent separation



naive MEM

- significant separation power lost



need proper modeling of detector transfer function
=> not yet conclusive / practical
but could even at this level be useful input to classifier

The Proposal - from fall 2021 - to today

Overview

- **event generation:**
  - start with only “ZHH” and “ZZH” (as usual each Z decay mode, separated into qq, ll, ee, vv)
  - 3 ECMs: 500 GeV, 550 GeV, 600 GeV
  - for aim ~ 2x the statistics Claude had (per ECM) => ~6 M events per ECM, i.e. 18M total

- **sim/rec:**
  - start with 500 GeV (for comparison to Claude approx. compatibility with IDR samples
  - eventually also 550 or 600 GeV, t.b.d. after generator-level (or SGV-level?) comparisons

- **other relevant SM backgrounds**
  => request later at least for one ECM, t.b.d. together with eg tt and ttH analysers:
  - 6f with at least 2 b’s, i.e. “ZZZ” / “**tt**” / “ZWW”
  - 4f with at least 2 b’s, i.e. “bbbb”, “vvbb”, “eebb”, “llbb” (can 2f, aa, ae, … be safely ignored?)
  - **ttH** / ttZ / ttg (g -> bb)
- possibly all with cut on m(bb) > 45 (?) GeV
  - Ok for ttH - but not useful for tt analyses?
  - Reduction in #evts worth it?

The Proposal - from fall 2021 - to today

Overview

- **event generation:**
  - start with only “ZHH” and “ZZH” (as usual each Z decay mode, separated into qq, ll, ee, vv)
  - 3 ECMs: 500 GeV, 550 GeV, 600 GeV
  - for aim ~ 2x the statistics Claude had (per ECM) => ~6 M events per ECM, i.e. 18M total

- **sim/rec:**
  - start with 500 GeV (for comparison to Claude approx. compatibility with IDR samples
  - eventually also 550 or 600 GeV, t.b.d. after generator-level (or SGV-level?) comparisons

- **other relevant SM backgrounds**
  => request later at least for ECM = 550 GeV together with eg tt and ttH analysers:
  - 6f with at least 2 b’s, i.e. “ZZZ” / “**tt**” / “ZWW”
  - 4f with at least 2 b’s, i.e. “bbbb”, “vvbb”, “eebb”, “llbb” (can 2f, aa, ae, … be safely ignored?)
  - **ttH** / ttZ / ttg (g -> bb)
- possibly all with cut on m(bb) > 45 (?) GeV
  - Ok for ttH - but not useful for tt analyses?
  - Reduction in #evts worth it?

The Proposal - from fall 2021 - to today

Overview

- **event generation:**
  - start with only “ZHH” and “ZZH” (as usual each Z decay mode, separated into qq, ll, ee, vv)
  - 3 ECMs: 500 GeV, 550 GeV, 600 GeV
  - for aim ~ 2x the statistics Claude had (per ECM) => ~6 M events per ECM, i.e. 18M total

- **sim/rec:**
  - start with 500 GeV (for comparison to Claude approx. compatibility with IDR samples
  - eventually also 550 or 600 GeV, t.b.d. after generator-level (or SGV-level?) comparisons

- **other relevant SM backgrounds**
  => request later at least for ECM = 550 GeV together with eg tt and ttH analysers:
  - 6f with at least 2 b’s, i.e. “ZZZ” / “**tt**” / “ZWW”
  - 4f with at least 2 b’s, i.e. “bbbb”, “vvbb”, “eebb”, “llbb” (can 2f, aa, ae, … be safely ignored?)
  - **ttH** / ttZ / ttg (g -> bb)
- possibly all with cut on m(bb) > 45 (?) GeV
  - Ok for ttH - but not useful for tt analyses?
  - Reduction in #evts worth it?

Our C3 guest members
(Caterina Vernieri and Dmitri Ntounis)
offer to contribute to the production
=> to discuss with Generator & Production Conveners: HOW

Next steps

Overview

- verify that 2f, aa, ea, … are irrelevant (based on IDR 500 GeV samples) -> Bryan
- train Dmitri on running Whizard in production system compatible way -> Mikael
- 4f / 6f: do-able with restrictions to at least on bbar pair? #evts with and without **invariant mass cut?** -> Mikael / Junping
- ttH / ttZ / tt(g->bb) do-able? -> Junping
- initially: pass all through SGV
- **full sim/rec to be decided together with production team**, depending on #evts, support from C3, priorisation in production system, …

Next steps

Overview

- verify that 2f, aa, ea, … are irrelevant (based on IDR 500 GeV samples) -> Bryan
- train Dmitri on running Whizard in production system compatible way -> Mikael
- 4f / 6f: do-able with restrictions to at least on bbar pair? #evts with and without **invariant mass cut?** -> Mikael / Junping
- ttH / ttZ / tt(g->bb) do-able? -> Junping
- initially: pass all through SGV
- **full sim/rec to be decided together with production team**, depending on #evts, support from C3, priorisation in production system, …

Outlook on mid-term schedule

- provide up-to-date 550 GeV self-coupling projection for EPPSU
  (person power: Bryan + Dimitri as of now, Julie from ~October - anybody else interested?)
- preview on analysis at ECFA WS in October
- status update at LCWS
- **obviously this is very tight, but we're not starting from zero, and this *is* important**