

Issues in the generated samples

Mikael Berggren¹

¹DESY, Hamburg

ILD swana, 26 Aug. 2026

CLUSTER OF EXCELLENCE
QUANTUM UNIVERSE



Introduction

Lately a few issues in the generated samples have been found. They affect both the 250 GeV full simulation mass-production, and the newer 550 GeV samples for the di-Higgs analysis.

- 1 Higgs to $\tau\tau$ in 'Higgs-less' samples
- 2 Wrong BRs in some 4f-hadronic samples
- 3 Rare corrupt input events
- 4 Wrong direction of compton-scattered electrons in GuineaPig

Higgs to $\tau\tau$ in 'Higgs-less' samples

- The issue

- Reported by Lea Kuttler:
- In the `4f_zz_l`, `4f_zz_sl` and `4f_zzorww_l` SM samples, $H \rightarrow \tau\tau$ events are present, although they shouldn't be - all Higgs channels should be in the dedicated samples, none in 'SM samples'

- The reason

- In the steering file to Whizard, the statement excluding the Higgs channel ("restrictions = !H") was in the wrong position, effectively making it a NOOP.
- NB. this *only* affects $H \rightarrow \tau\tau$, since that on one hand at Whizard level all quark-masses are set to zero (i.e. no coupling to the H), and on the other hand the 6f samples were done avoiding Higgs channels in a different way, not affected by the error, so no $H \rightarrow ZZ^*$ or WW^* .

Higgs to $\tau\tau$ in 'Higgs-less' samples

• The issue

- Reported by Lea Kuttler:
- In the `4f_zz_l`, `4f_zz_sl` and `4f_zzorww_l` SM samples, $H \rightarrow \tau\tau$ events are present, although they shouldn't be - all Higgs channels should be in the dedicated samples, none in 'SM samples'

• The reason

- In the steering file to Whizard, the statement excluding the Higgs channel ("`$restrictions = !H`") was in the wrong position, effectively making it a NOOP.
- NB. this *only* affects $H \rightarrow \tau\tau$, since that on one hand at Whizard level all quark-masses are set to zero (i.e. no coupling to the H), and on the other hand the 6f samples were done avoiding Higgs channels in a different way, not affected by the error, so no $H \rightarrow ZZ^*$ or WW^* .

Higgs to $\tau\tau$ in 'Higgs-less' samples

- The solution/mitigation

- Reject events in those three samples where the true $|M_{\tau\tau} - 125| < 50 \text{ MeV}$
- Look for the *first* occurrence of τ :s in MCParticles - this would be the pre-FSR ones.
- Scaling factors should not need to be changed: cross-section and number of (used) events in the sample decreased by the same factor.
- Side remark: The Higgs *is* part of the SM: Also in non-Higgs analyses, don't forget to include the Higgs samples (the inclusive ones under dst-merged/250-SetA/higgs (nothiggs_excl)) in the background !

Higgs to $\tau\tau$ in 'Higgs-less' samples

- The solution/mitigation

- Reject events in those three samples where the true $|M_{\tau\tau} - 125| < 50$ MeV
- Look for the *first* occurrence of τ :s in MCParticles - this would be the pre-FSR ones.
- Scaling factors should not need to be changed: cross-section and number of (used) events in the sample decreased by the same factor.
- Side remark: The Higgs *is* part of the SM: Also in non-Higgs analyses, don't forget to include the Higgs samples (the inclusive ones under dst-merged/250-SetA/higgs (nothiggs_excl)) in the background !

Wrong BRs in some 4f-hadronic samples

- The issue

- Reported by Bryan Bliewert:
- In `4f_zz_h`, the total cross-section is OK, but the relative Br(4 same quarks) and Br(2 one type, 2 another) are wrong, if resonance insertions are requested in the Whizard steering.
- Resonance insertions are required for getting the di-jet masses correct.

- The reason

- Bug in Whizard.
- Found and fix known (by Claude Code...). Reported, but not yet implemented.

Wrong BRs in some 4f-hadronic samples

- The issue

- Reported by Bryan Bliewert:
- In `4f_zz_h`, the total cross-section is OK, but the relative Br(4 same quarks) and Br(2 one type, 2 another) are wrong, if resonance insertions are requested in the Whizard steering.
- Resonance insertions are required for getting the di-jet masses correct.

- The reason

- Bug in Whizard.
- Found and fix known (by Claude Code...). Reported, but not yet implemented.

Wrong BRs in some 4f-hadronic samples

- The solution/mitigation
 - Reweighting would be complicated (impossible?):
 - How much wrong the weight of a given event is depends both of the quark composition, and how many resonance insertions were done.
 - So: Best redo these channels once the fix is verified and implemented by the Whizard authors.
 - Would be 10.2 Mevents in 11 generator jobs, producing 102 files to re-simulate and reconstruct.

Wrong BRs in some 4f-hadronic samples

- The solution/mitigation
 - Reweighting would be complicated (impossible?):
 - How much wrong the weight of a given event is depends both of the quark composition, and how many resonance insertions were done.
 - So: Best redo these channels once the fix is verified and implemented by the Whizard authors.
 - Would be 10.2 Mevents in 11 generator jobs, producing 102 files to re-simulate and reconstruct.

Rare corrupt input events

● The issue

- Bryan Bliewert reported occasional crashes in SGV, underlying reason identified by MB:
- Sometimes total E & p not conserved, and/or daughters of decayed particles are missing.
- Makes SGV crash (but not FullSim, even though it should ...)
- Very rare (sub-permil)

● The reason

- Unclear - has to do with the Whizard/Pythia interface. Occurs if the event is not fully generated (e.g. some decay products are not generated, but the event is still considered as a valid event).
- The problem is that the Whizard/Pythia interface does not check if the event is fully generated before it is passed to the SGV.
- The problem is that the Whizard/Pythia interface does not check if the event is fully generated before it is passed to the SGV.

Rare corrupt input events

- The issue

- Bryan Bliewert reported occasional crashes in SGV, underlying reason identified by MB:
- Sometimes total E & p not conserved, and/or daughters of decayed particles are missing.
- Makes SGV crash (but not FullSim, even though it should ...)
- Very rare (sub-permil)

- The reason

- Unclear - has to do with the Whizard/Pythia interface. Occurs if
 - ISR on ('ISR on' inevitably creates some tiny E & p non-conservation, but not of the size seen)
 - Resonance insertions on
 - Pythia's MSTP(125) set to 2 : Save the full details of the parton shower, needed for TrueJet to work.

Rare corrupt input events

- The solution/mitigation
 - Mitigation: In **SGV**, the consistency of the input is now verified (E & p consistent within 1/1000, no missing daughters, correct charge). Event skipped (and reported) if inconsistent.
 - Should similar check to be included in FullSim ?
 - Solution: ???

Wrong direction of compton-scattered electrons in GuineaPig

- The issue

- Reported by Shriyansh Ranjan:
- The direction of compton-scattered electrons and poistrons are exchanged in GuineaPig.
- Has no implication for energy-symmetric collisions, but is relevant for HALHF.

- The reason

- Bug in GuineaPig

- The solution/mitigation

- Fix the bug
- No action needed for non-HALHF simulation

Wrong direction of compton-scattered electrons in GuineaPig

- The issue

- Reported by Shriyansh Ranjan:
- The direction of compton-scattered electrons and poistrans are exchanged in GuineaPig.
- Has no implication for energy-symmetric collisions, but is relevant for HALHF.

- The reason

- Bug in GuineaPig

- The solution/mitigation

- Fix the bug
- No action needed for non-HALHF simulation

Wrong direction of compton-scattered electrons in GuineaPig

- The issue

- Reported by Shriyansh Ranjan:
- The direction of compton-scattered electrons and poistrans are exchanged in GuineaPig.
- Has no implication for energy-symmetric collisions, but is relevant for HALHF.

- The reason

- Bug in GuineaPig

- The solution/mitigation

- Fix the bug
- No action needed for non-HALHF simulation

Summary

- Three issues found with the generated samples.
- Affects the 250 GeV mass-production, so action is needed:
 - There are $H \rightarrow \tau\tau$ in the 'Higgs-less' samples $\Rightarrow$ Remove events with true $\tau\tau$ mass within 50 MeV of the Higgs mass from analysis.
 - In samples containing both four quarks of the same flavour, and two of one flavour, two of another, the relative fractions are wrong $\Rightarrow$ Need to re-do these samples.
 - There is a tiny risk that energy is not conserved in some events, and/or that some final state particles are missing $\Rightarrow$ In SGV, use the latest version. In FullSim: Ignore, or include a consistency check function?
- In addition, an issue was found in GuineaPig, but it will only affect future HALHF studies.

Summary

- Three issues found with the generated samples.
- Affects the 250 GeV mass-production, so action is needed:
 - There are $H \rightarrow \tau\tau$ in the 'Higgs-less' samples $\Rightarrow$ Remove events with true $\tau\tau$ mass within 50 MeV of the Higgs mass from analysis.
 - In samples containg both four quarks of the same flavour, and two of one flavour, two of another, the relative fractions are wrong $\Rightarrow$ Need to re-do these samples.
 - There is a tiny risk that energy is not conserved in some events, and/or that some final state particles are missing $\Rightarrow$ In SGV, use the latest version. In FullSim: Ignore, or include a consistency check function?
- In addition, an issue was found in GuineaPig, but it will only affect future HALHF studies.

Summary

- Three issues found with the generated samples.
- Affects the 250 GeV mass-production, so action is needed:
 - There are $H \rightarrow \tau\tau$ in the 'Higgs-less' samples $\Rightarrow$ Remove events with true $\tau\tau$ mass within 50 MeV of the Higgs mass from analysis.
 - In samples containg both four quarks of the same flavour, and two of one flavour, two of another, the relative fractions are wrong $\Rightarrow$ Need to re-do these samples.
 - There is a tiny risk that energy is not conserved in some events, and/or that some final state particles are missing $\Rightarrow$ In SGV, use the latest version. In FullSim: Ignore, or include a consistency check function?
- In addition, an issue was found in GuineaPig, but it will only affect future HALHF studies.