# Weekly report

Yu Zhang

IHEP

May 22, 2017

Table 9: About the Wh column, the last two rows. Are you sure about the drop from 14.2% to 0.533% from the next-to-last to the last row in the table? The W decays about 23% of the time to either electron or muon. Now 14.2% * 0.23 = 3.3%. And 0.533/3.3 = 0.16. This seems VERY low to have an efficiency for a single electron or muon of only 16%, even including the acceptance. Similar arguments for the Zh and tth numbers: For Zh, the fraction of Z bosons that decay to ee or mumu are 6.7%. And there, you even have two possible leptons to select from. So 12.9% * 0.067 = 0.86%. Then, 0.354/0.86 = 41%. This seems actually a bit more reasonable. For tth, we have roughly $1\text{-}0.8^2 = 0.36$ of the events where at least one W decays into an electron or muon. So 6.18% * 0.36 = 2.2%. Then, 1.89/2.2 = 0.86. Also this seems reasonable. So why should we have for tth a lepton efficiency of 86% while for Wh, we have only 16%. Am I missing something? Did I make a mistake in the estimations above?

- I think this calculation is correct. We need to check the number.

Section 7.3, 7.3.1, 7.3.2:

It would be good to also show the fitted parameters, including their uncertainties, on each plot (or in a Table).

Fig. 12: It should be stated in the caption that the fit is performed in (a) and the resulting funtion is simply drawn in (b), using the normalization from (b) or fitting the normalization in (b)?

Also, while it is nice to see that the shape determined from 0-lep does somewhat work for 1-lep, I feel that one should use a quantification of the degree of non-matching as a systematic nuisance parameter in the final fit. For example the uncertainties/spreads of the individual parameters of the ExpPoly2 function between the various 0-lep and corresponding 1-lep regions could be used for this. Surely, the comparission in Fig. 12 gives sufficient statistics for this, possibly also the ones of Fig. 13. That is also one reason why I am asking for the post-fit parameters, including uncertainties, for all these plots. One problem will of course be that the individual parameters are not going to be uncorrelated.

- need to show the parameters
- need to modify Fig 12
- introduce the nuisance parameters for bkg shape?

On line 309, you say that both the background shape and normalization are determined simultaneously in the fit to SR and CR. In the above comment, I have asked about how you treat systematic uncertainties regarding the shape differences. How do you treat systematic differences regarding the normalization? ... Ah, I just see that this is explained in Section 8.4, good.

Why is the cas of only RevID not shown in Fig. 13?

- We need to consider the uncertainty of background shape consistency between CR and SR.
- Add RevID in fig 13.
- @Qi

Section 7.3.3:
Fig. 14: It seems clear to me that the shape doesn't fit in the signal region plots on the left. The line has a different slope compared to the RevData+MC points. This connects again to my worry about constraining the shape parameters to be completely identical between the SR and CR. You still need to define $N_{ss}$ in the text. It is still only implicit. Please also unify the names in Table 13, e.g., in Table 13, you use $S_{ref}$ and DeltaS, while in the text, you use $DeltaN_{ss}$ and $N_{S.exp}$.
What is the conclusion? Is the spurious signal test successfully passed? Is the maximum $N_{ss}$ added as an uncertainty or so (I just see that this is discussed in Sec. 8.4. Maybe add a forward reference here?)?

- My work. Surely the slope is different in fit and data point without pTyy cut, we all know it, but it can pass spurious signal criteria.
- make the text consistent. add a conclusion.

Section 8.1:
To which parts of the fit do you assign the luminosity uncertainty to? As you fit your background normalization from a data control region, the luminosity uncertainty on that part drops out. I would thus argue that this uncertainty should only be attached to the SM single higgs-boson signal MC and the hh signal MCs.

- Right

Section 8.2:
Lines 373-376: You can and should apply the same uncertainty also on Zh. Something of similar order is probably also needed for ggh, but since that expected yield is so low, it probably won't change much.
Actually, does the Pythia8 Zh sample that you use include the gg-¿Zh contribution?

- Is it easy for us to have Zh theo sys?
- gg-¿Zh is not included.

Section 8.3:
Table 18: The pileup reweighting uncertainty for ggh seems excessively large with 11%, compared to the others. Maybe an issue of limited MC statistics and/or some high-weight events?
Section 9:
Section 1.2 says that I shouldn't read this yet...

- @Qi could you check the pile up weight?

# Follow-up about previous comments

l128: How is this reweighting done?
¿ It is implemented in the package by theorists and more details can be retrieved from ref.[13] in note version 0.2
Thanks. Please also put the citation here. Otherwise, it is not clear.

I134: Are the W bosons allowed to decay into anything? Or is one W boson restricted to decay into leptons and the other into hadrons?
¿ W decays inclusively
Please state this in the supporting note.

Fig. 1: At which cut level are these (and all other plots) shown? Please always specify this in the captions. Also, are we looking here at truth or reco quantities? The text seems to indicate that this are truth quantities. But why are there then dips at —eta— 1.5 in (d)?
¿ These are in truth level. There are —eta— and pT cuts in truth level in the HGam framework.
Thanks. Please always specify in the caption which cuts are done, as this is apparently not right after the generator.

Table 4: Please add a column with the cross sections that you assumed for these samples, including at which order that cross section is calculated.
¿ Yes we can add xs, though they are not precise from theoretical side and are not used in the analysis
Now, you use a jjjyy sample. For the cross-section, you have now a "*". What does this mean? Please either clarify in the caption or enter the actual value.

- Actually, we always don't care about the cross section since we normalize the MC to data sideband.

l174-178: Is this really needed? How often does this find another primary vertex compared to the ljj vertex?
¿ This is the standard HGam recommendation. We do not have a ljj customised version.
True. But how often do you find a different vertex in your signal sample(s) with the gammagamma pointing w.r.t. to the "standard" primary vertex, i.e., sum $pt^2$?

- My work, asked by Karsten, Giovanni and Chris.

l192+l194: "w.r.t. the primary vertex": which one? The original one or the one you describe in l174-178?
¿ In contact with experts about this.
Do you have the answer by now?

- w.r.t. NN vertex

¿ We followed HGam standard overlap removal procedure. The overlap is based on photons with looseID and no isolation requirement. antiID photons mean failing tightID but still passing looseID. Thus, antiID and antiIso photo events have been properly done with overlap removal and consistently with ID+ISO photon events as well.
Please write down the precise procedure in this section. As it seems to be a non-standard overlap removal procedure, has this been presented in an ASG meeting and blessed? I believe this is the guideline for such non-standard cases (maybe even required?).

- add to the note
- of course, it should already be presented.

I221: What is the overall trigger efficiency for this trigger for signal MC in the signal regions?
¿ around 70
No. Table 5 shows the trigger efficiency before most offline cuts are applied. One ususally wants to go to the signal region and as a last cut apply the trigger selection. That will give you the trigger efficiency.

- It can be easily got.

l243: Why not exactly one lepton but rather at least? Fig. 5(b) seems to indicate that this would be beneficial. How many events of the type Z(-¿ee, reconstructed as gamma,gamma)W(-¿l,nu)jj survive your signal selection befor cutting on m(gamma,gamma)¿110 GeV?
¿ Good point. We probably need to consider using ==1 lepton in the next round (version ¿= 0.4).
I do remember seeing some information about that in one of the recent EB meetings. Maybe add that information to the supporting note?

- add this to the supporting note.

¿ This table aims at showing the overall uncertainties from different objects. Each number is a group of uncertainties from many eigen vectors or details uncertainties. So quadratic sum up is used conservatively. Then we lose the sign. Nevertheless, appendices records detailed rate uncertainties down to NP level. This table will significantly updated in the next round (version ¿= 0.4).
Please state in the captions that each line is the quadratic sum of their respective individual components.

How exactly is the 0-lepton control region used in the fit?
¿ Simultaneous fit with 0-lep and 1-lep control region and in both regions the background modelling is shared
Do you not have any kind of systematic associated with possible shape differences between the two regions? I guess I will have to wait for a full description of this in Section 9, once it is there.

l402: "The semi-leptonic and hadronic channels are orthogonal by definition" Is this still true for the 0-lepton control region?
¿ yes, we require 0lep with ¿= 2 jets, since fully hadronic channel is not in anymore.
I guess this is OK. But the fully-hadronic analysis is not abandoned. It is just postponed. Without having some explicit orthogonalization criteria, it will not be possible to combine the two results once the fully hadronic channel is ready.