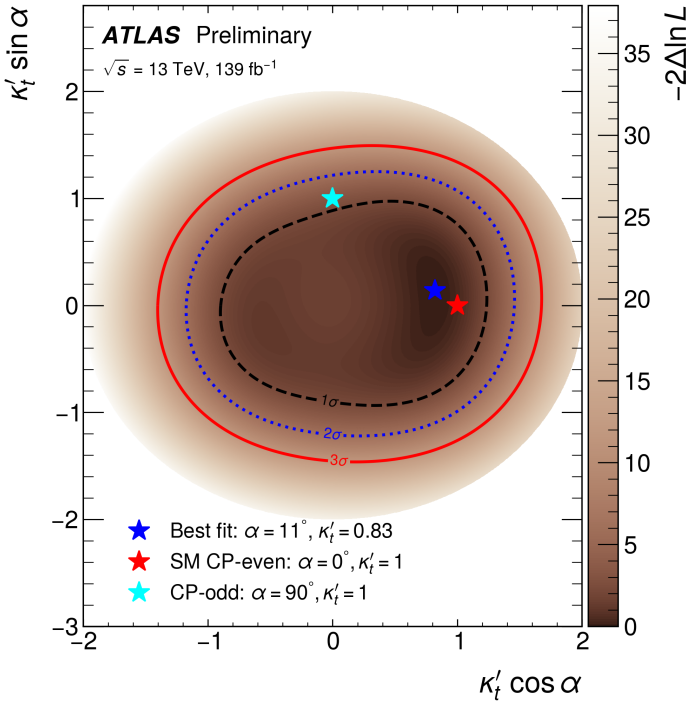
ttH(bb) CP Paper – Editors Report



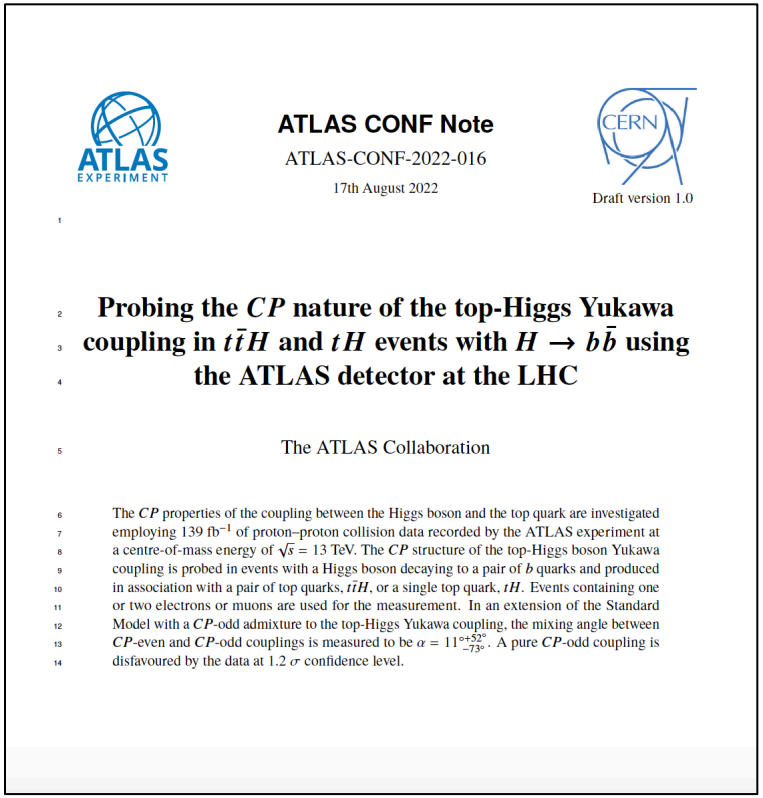
Brian Le, Yang Qin, R. Gonçalo

Editorial Board meeting 30/8/2022

At a glance

Reference links:

- Last presentation in HTOP meeting: https://indico.cern.ch/event/1190122/
- CDS record:https://cds.cern.ch/record/2802607
- New version (17 Aug., not yet in CDS): https://cernbox.cern.ch/index.php/s/AGC6qOzs3Al2hop
- Diff to previous version: https://cernbox.cern.ch/index.php/s/QpIk8NkK7gVM2Jt
- Comments and answers collected in Google Doc here:



ATLAS CONF Note

ATLAS-CONF-2022-016

17th August 2022

Draft version 1.0

Probing the CP nature of the top-Higgs Yukawa coupling in $t\bar{t}H$ and tH events with $H \to b\bar{b}$ using the ATLAS detector at the LHC

The ATLAS Collaboration

The CP properties of the coupling between the Higgs boson and the top quark are investigated employing 139 fb^{-1} of proton–proton collision data recorded by the ATLAS experiment at a centre-of-mass energy of $\sqrt{s}$ = 13 TeV. The CP structure of the top-Higgs boson Yukawa coupling is probed in events with a Higgs boson decaying to a pair of b quarks and produced in association with a pair of top quarks, $t\bar{t}H$, or a single top quark, tH. Events containing one or two electrons or muons are used for the measurement. In an extension of the Standard Model with a CP-odd admixture to the top-Higgs Yukawa coupling, the mixing angle between CP-even and CP-odd couplings is measured to be $\alpha = 11^{\circ}{}^{+52^{\circ}}_{-73^{\circ}}$. A pure CP-odd coupling is disfavoured by the data at 1.2 σ confidence level.

News (see Zak's talk):

- New post-fit uncertainties from xRooFit (TRexFitter not propagating scale factors)
- New tH cross sections and uncettainties from Yellow Report
- **New yield tables**
- As far as we can see: **all comments addressed**

Questions & Answers

- **[Haichen]** SL4: According to Top physics experts, 4FS is better for top modeling. There is also a difference in the scale choices between the MC calculation and the YR. Not completely clear that the effect is small, e.g. in the boosted SR. There was a preference in the meeting to harmonize to the YR numbers, which also come with uncertainties that in principle should cover the 4FS vs 5FS differences. Agreed to revisit for Draft2.
- > We have now moved to the YR numbers and there has been little difference in the overall result.

Questions & Answers

- **[Fabio]**: SL9 Is this quantity defined for tH events?
- Answer: The point is all events are reconstructed assuming a ttH model. In tH events, one of the tops is therefore fake. This will be clarified in Draft2.
- > We got conflicting comments on whether or not we should explain that tH is reconstructed with a fake top. It appears that some people think this is obvious and don't need explanation. We now have a statement in the introduction saying that the analysis is optimised for ttH. It should follow then that the reconstruction is based also on ttH.

Also:

- **[Giacinto]** L49-L50: something which is unclear at this point and later; even if you now consider tH to be signal, is there any change you apply to the classification and reconstruction BDTs? In other words, does the analysis also explicitly target tH as signal? If not, it may be good to mention that here. I.e. that although tH is not targeted explicitly, it is considered as signal in this analysis, for the reasons mentioned above.
- > Already changed following other comments: we explicitly say that the analysis is optimised based on ttH only

And

- **[Dresden]** L71 When discussing the final states as (decay) channels, it might be worth hinting at the interplay with the fact that you are assuming both ttH and tH as your signal, e.g. dilepton would only be relevant for the former I guess.
- > Our analysis was actually optimised for ttH, tH is only included for completeness. In fact dilepton is not sensitive to tHjb, only to tWH.

Questions & Answers

- **[SRs and CRs]** -I find the introduction of a first series of categories including the PSRs (preliminary segnal regions), and then a second iteration introducing how the PSRs are further distinguished into other categories a bit confusing. Also, in the first step, the subdivision between SRs and CRs based on S/B seems very arbitrary, and it would be better to just distinguish between regions where a full reconstruction of the event is possible, and where it is not. So I believe a better way to structure this part of the paper would be to first describe all the ingredients, i.e. physics objects based on which the selection is defined, then introduce the categories (>=6 and >=4 jets) where a full event reconstruction is possible, describe briefly reconstruction and classification BDTs, and the CP-sensitive observables, and only then move to describe all the analysis categories in one go. You could still keep most of Table 1 and Table 2, by just focusing in Table 1 on all categories where a full event reconstruction is not possible (so all regions except for the PSRs), and keep what you have in Table 2, but having the text describing all categories in one go. This should also make the flow of the text better in several places, in addition to removing the problem of defining the PSRs (preliminary signal regions) first and then redefining them again later on. ==>Notice: I would be completely fine not to do this for the CONF note conversion but only for Draft 2, since it requires some significant changes/reshuffling of the content.
- > We have made more of an effort to explain why there is this two step procedure. To address your first comment the first step, the division between CRs and what we called PSRs is aimed specifically at replicating the CR structure of the STXS analysis. The CRs in this step are solely for controlling the different background components (due to the number of jets). What we call CRs in the second step are really just signal-depleted regions. We have removed mention of the signal purity in order to avoid confusion as to the purpose of each of the regions.
- In your suggestion we would define the 6j and 4j region and discuss the tools we use but then come back to the analysis categorisation at which stage we are then simultaneously discussing CRs designed for background constraints and also CRs which are residuals of our classification BDT split. We believe that mixing the two ideas is not ideal.
- Our compromise is to highlight these PSRs are actually regions for **training** our classifiers by denoting them as **training regions (TRs)**. We hope by shifting the focus, this will clarify the distinction between the two steps.

Questions & Answers

- **[ttH and tH model]** L134-146: I don t see the overall cross-sections used for the ttH and tH samples mentioned anywhere here. Do you use the cross-sections straight out of the MCs? If not, could you mention what values you use, and where these comes from? And how to do you handle the non-SM case? (same K-factor?)
- > The ttH cross-sections are indeed taken from the Yellow report and a consistent k-factor is used for both SM and non-SM signals. For tH the Yellow report seems to have recommendations for the 5FS and not the 4FS. We have checked with the current tH team and they confirm they take the cross-sections from the MC.
- We have added in the draft the cross sections we use for ttH and tH, and that the k-factor derived from the SM case is applied to all CP scenarios.
- From YR, tH cross sections are shown for tHjb t-channel and s-channel separately:
  - 74.26 fb and 2.875 fb – although the contribution from interference is not clear.
  - For tWH, it is 15.17 fb.
- Cross section from MC that we use:
  - tHjb: 60.1 fb
  - tWH 16.7 fb
- There's quite a big difference in tHjb. But given the different FS this is probably not so surprising. The impact ultimately was found to be very small.

Questions & Answers

- **[5FS to 4FS uncert.]** L252: can you justify why you think this procedure is reasonable? can you say anything about how close the tt+>=1 b-jet model is for the nominal 4flavor tt+bb simulation and the 5flavor tt+jets simulations? if these are sufficiently close, then you may expect the approximation you make to hold?
- > Indeed this is only an approximation but unfortunately we can only evaluate systematics based on samples that exist. The 4FS ttbb sample is a very special case with dedicated implementations. The code was provided by the authors with a special release to provide us the prediction early on - see ref 53. And we only had this one sample, Powheg+Pythia8, no alternative sample generated with e.g. 4FS amcnlo with Pythia, or Herwig. But this sample is really the state-of-the-art. It represents our best knowledge of this background and this is the main reason for using it.
- The tt+bb 4FS and 5FS difference isn't small - but this difference being small isn't necessarily a good justification. However, we do have a dedicated 4FS vs. 5FS systematics as described later which was necessary to cover mismodelling observed in our analysis. This uncertainty is expected to cover any potential missing effect associated with the procedure to evaluate the systematics we are using.
- STXS analysis used a similar recipe, even without the 5FS vs. 4FS systematic.

Questions & Answers

- **[Michigan]** Line 186: Naively, these numbers seem low (one might assume that with 4 b-jets, there are 3 ways of pairing them, so even guessing randomly would be a 33% correct assignment fractions). Is this because sometimes there are b-jets aside from the 2 H b-jets and 2 t b-jets and/or missing b-jets from the H or t decays, etc., in which case even a perfect BDT wouldn t have 100% correct assignment rate? If so, are these rates relative to the total number of events in PSR, or relative to the total number of events in which the b-jets from the tops and H s are correctly reconstructed and b-tagged?
- > Assuming perfect b-tagging (100% efficiency and 0 fake), if we have 4 b-jets and pick two out of them, there are 6 possible combinations. So a random pairing would result in ~17% instead of 33%. Adding effect of b-tagging this number will further reduce.

- But you're right that these fractions are calculated wrt to the total number of events in the PSR (now renamed to TR), for each region where a reconstruction BDT is trained. Indeed there're quite some events don't contain the jets that we need to start with, so these events will not be able to count in the numerator of the efficiency. See figure 195 of the previous internal note for tthbb https://cds.cern.ch/record/2244360/files/ATL-COM-PHYS-2017-079.pdf

Questions & Answers

- **[Athens]** L 143: could you add the diagrams for single top H associated production
- > If we were to include diagrams, we will need to include all of them (ttH, tHjb, and tWH with t-H coupling and W-H coupling), otherwise it becomes weird. This means we need at least 4 diagrams to show the most representative processes. There are plenty of references that we quote in this paper include these diagrams. So we propose to not include the diagrams at all.