Imperial College London

Faculty of Natural Sciences

February Report

PhD Journal: Year 1

Author:
Jonathon Langford

Supervisor:
Dr. Nick Wardle

6th April, 2018

1 Trilinear analysis

An analysis to investigate the possibility of trilinear Higgs coupling determination via single-Higgs differential measurements. Follows work of Maltoni et. al [1, 2].

1.1 Introduction and current position

Introduction and current stage of analysis.

1.2 Theory

Trilinear theory section

1.3 Week beginning 02/04/2018

Use following link to see plots related to this weeks work.

1.3.1 Deficit at low $p_T(\gamma\gamma)$

- flashgg::ZHLeptonicTag sample from running ZHLeptonicTag dumper on inclusive VHTOGG LHC samples[1].
- Madgraph sample refers to the sample generated using code provided by trilinear authors. Events have been through the Pythia8 and Delphes, using CMS card.
- **Importance of this work...**
  - Understand and get to grips with flashgg framework.
  - Investigate ZHLeptonicTag, could maybe give insight to $H\gamma\gamma$ analysis, and offer improvements.

Observe deficit of flashgg::ZHLeptonicTag events in low gen-level $p_T(\gamma\gamma)$ region, compared to madgraph event sample (figure 1). Inclusive $p_T(H)$ spectrum from full VHTOGG LHC sample matches shape of madgraph spectrum, so there is no significant difference in intrinsic kinematics. Suggests that the shift in distribution occurs at selection level. Note, at this stage the madgraph event sample has limited selection on events ($p_T(\gamma)$, ΔR, $p_T(\ell)$, $m_{\ell\ell}$ in Z window).

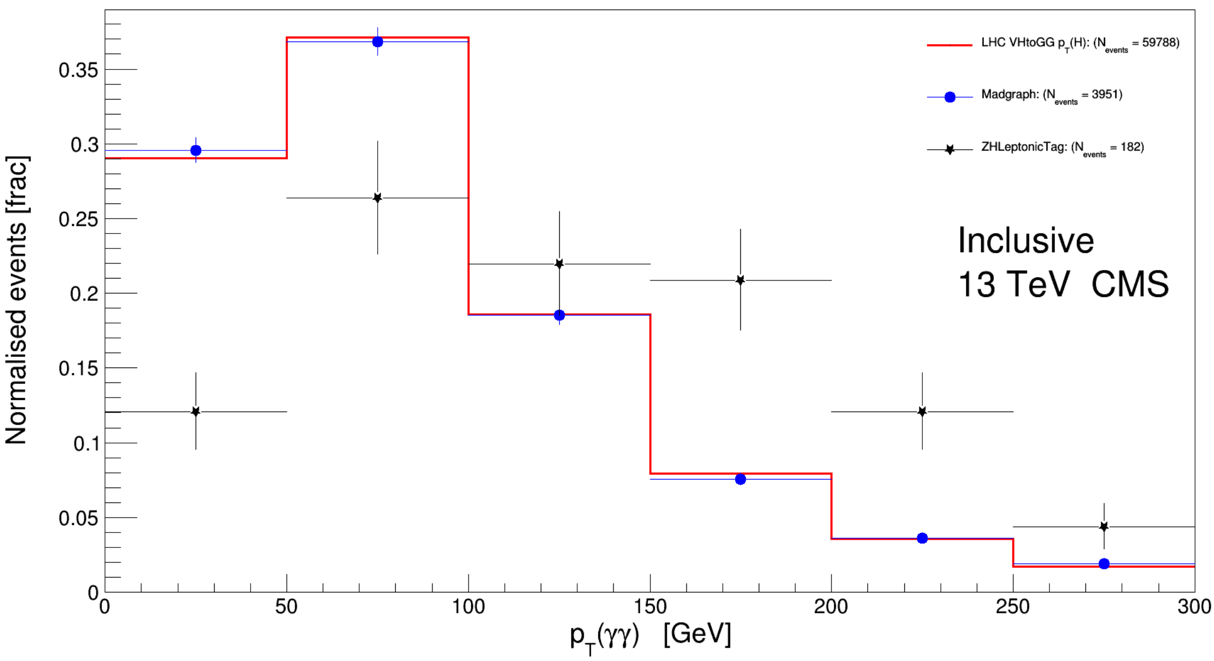
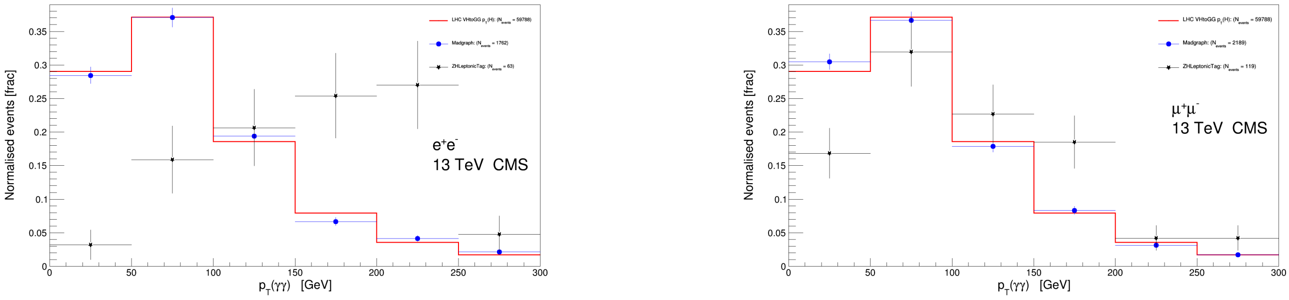


Figure 1: Comparing the $p_T(\gamma\gamma)$ spectrum of madgraph event sample to flashgg::ZHLeptonicTag events (from VHTOGG LHC sample). $N_{events} = 182$.

- Split events according to lepton flavour in Z decay: $Z \to e^+e^-$ and $Z \to \mu^+\mu^-$, and plotted resulting $p_T(\gamma\gamma)$ spectra (figures 2a & 2b). See effect significantly more pronounced in ee type events. Note higher eff $\times$ acc for $\mu\mu$ type events, as expected.
- Introduced cut-flow, with output in endJob() method of ZHLeptonicTagProducer showing number of events and mean $p_T(H)$ at each cut. Results shown in screenshot in figure 3. Key points are as follows:
  - Diphoton selection has small effect on mean $p_T(H)$ (vertex). Not causing the discrepancy in $p_T(\gamma\gamma)$ spectrum.
  - See huge leap in $p_T(H)$ after lepton selection. Trialled two different $p_T(\ell)$ requirements at 20 and 10 GeV, small increase in acceptance in first bin of $p_T(\gamma\gamma)$ spectrum for $\mu\mu$ type events, but not

[1]Sample: VHToGG_M125_13TeV_amcatnloFXFX_madspin_pythia8,
Campaign: RunIISummer16-2_4_1-25ns_Moriond17.



(a) $Z \to e^+e^-$, $N_{events} = 119$ (b) $Z \to \mu^+\mu^-$, $N_{events} = 63$

Figure 2: Comparing the $p_T(\gamma\gamma)$ spectrum of madgraph event sample to flashgg::ZHLeptonicTag events (from VHTOGG LHC sample), split into ee and $\mu\mu$ type categories.

enough to describe deficit. Important to investigate lepton selection step further (deconvolve).

```
##########################################################
                      CUT COUNTERS

00) Total events: 59788
00) Total mean pT(H): 92.9647

01) (Vertex) Events: 29843
01) (Vertex) pT(H): 100.806

02) (LeadPhoPToverMass) Events: 29084
02) (LeadPhoPToverMass) pT(H): 102.344

03) (SubleadPhoPToverMass) Events: 25092
03) (SubleadPhoPToverMass) pT(H): 100.952

04) (Photon ID) Events: 25092
04) (Photon ID) pT(H): 100.952

05) (Diphoton MVA) Events: 23186
05) (Diphoton MVA) pT(H): 103.895

07a) (Lepton selection (pT20,eta,vtx,ID,dR)) Events: 188
07a) (Lepton selection (pT20,eta,vtx,ID,dR)) pT(H): 137
      07a) Electron events: 64
      07a) Electron pT(H): 174.159
      07a) Muon events: 124
      07a) Muon pT(H): 117.822
07b) (Lepton selection (pT10,eta,vtx,ID,dR)) Events: 214
07b) (Lepton selection (pT10,eta,vtx,ID,dR)) pT(H): 128.793
      07a) Electron events: 67
      07a) Electron pT(H): 169.704
      07a) Muon events: 147
      07a) Muon pT(H): 110.147

08) FINAL (inc (Zmass window)) Events: 182
08) FINAL (inc (Zmass window)) pT(H): 136.553
      08) Electron events: 63
      08) Electron pT(H): 172.739
      08) Muon events: 119
      08) Muon pT(H): 117.395

##########################################################
```

Figure 3: Screenshot of ZHLeptonicTag cut-flow, showing number of events and mean gen-level $p_T(H)$ at each cut.

1.3.2 ttH Event generation

Successfully generated ttH events in madgraph without segmentation fault (authors updated code). Run through PYTHIA8 with Higgs decay limited to $H \to \gamma\gamma$ in pythia config. See link for Feynman diagrams of LO and $\mathcal{O}(\lambda_3)$ ttH processes.

- Important as ttH has largest intrinsic C_1 value of all dominant Higgs production modes. Thus, increased effect of including electroweak loops on kinematics distributions, and thus increased sensitivity to κ_λ.
- Figure 4 demonstrates can reach as high as 9% in differential C_1 in some bins of kinematic distribution. Compare to ZHLeptonic with max 5%.

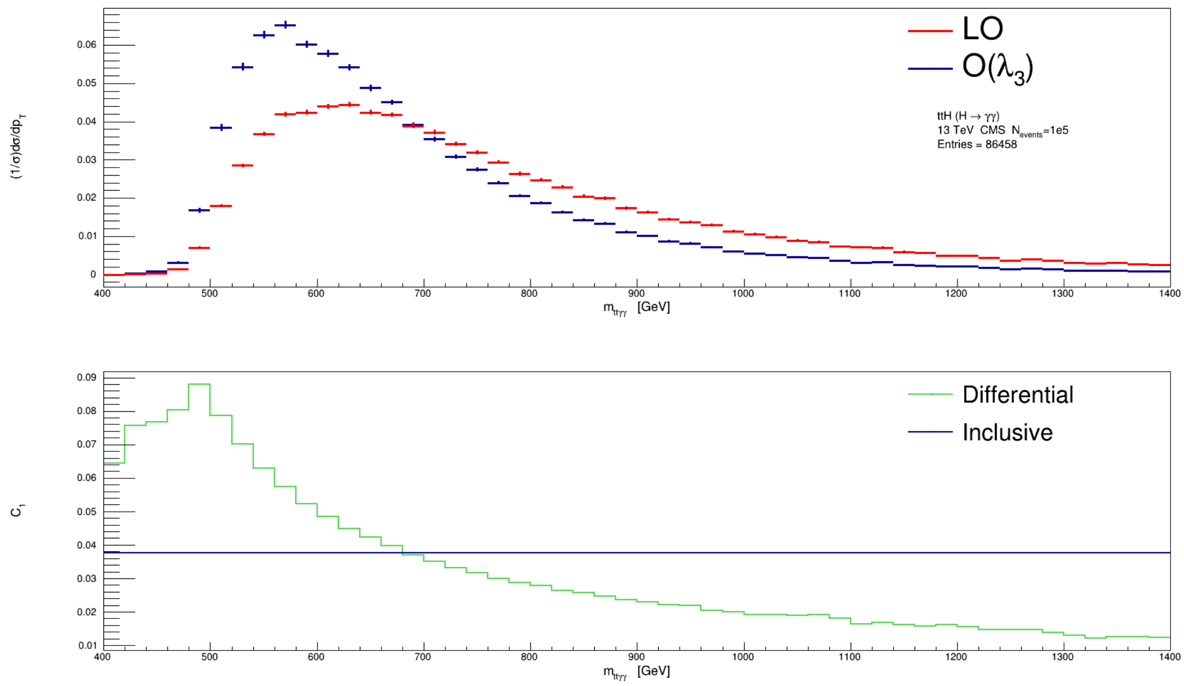


Figure 4: Gen-level $m_{tt\gamma\gamma}$ distribution, with C_1 value in each bin shown below. Obviously, would use much coarser binning when it comes to setting up analysis as statistically limited.

1.3.3 Setting up analysis

- Want to set up analysis in both ZHLeptonic and ttH(inclusive), as these production modes offer the two largest intrinsic C_1 values. Structure of analysis as follows:
  - Calculate C_1 in bins of some gen-level kinematic distributions (use coarse binning as will be statistically limited). At some point will need to optimise bin structure to be used throughout analysis.
  - Reconstruct event (either directly in samples we will generate or via flashgg:Tags). Plot same kinematic distribution at reco-level to see event migration between bins.
  - Plot diphoton invariant mass spectrum in each bin (i) of kinematic distribution and extract signal strength: μ_i, where $\mu_i(C_1, \kappa_\lambda)$.
  - Fit using likelihood to extract κ_λ given C_1 values and taking into account event migration between bins. Asimov dataset at 3000 fb^{-1}, measure sensitivity?

- **ZHLeptonic**
  - Variables which seem to offer largest differential C_1 values: $p_T(\gamma\gamma)$ and m_{ZH}. Mirrors results of [1].
  - Issue of low MC statistics given BF and eff $\times$ acceptance.
- **ttH(inclusive)**
  - Variables: $p_T(\gamma\gamma)$ and m_{ttH} + ...
  - Reconstruction: extremely difficult to reconstruct to top quarks in an event, therefore variables like m_{ttH} may be difficult at reco level.
  - Have to consider ttHHadronic and ttHLeptonic separately and then combine at likelihood level. Investigate how top reconstruction performed in flashgg or by top quark group in differential measurements.
    * ttHHadronic: tag requires 4 jets, with atleast one b-tagged. Therefore may only be able to reconstruct one top to ensure reasonable acceptance. Choose 3 jets which give mass closest to m_t?
    * ttHLeptonic: more difficult due to presence of ν. Use variable like visible mass? Kinematic fitting methods?

1.3.4 Imperial Hgg Meeting discussion

- HL-LHC Delphes events do not seem to be coming any time soon. Perform event generation ourselves including signal + background. Seth will ask about High Lumi Cards.
- Introduce kinematic cuts into madgraph events to mimic flashgg::Tags, seeing effect on acceptance. e.g b-tagging, η. Should these be on gen-level or reco-level?
  - Interesting point: η cut in lepton selection limiting to barrel thus picking out high $p_T(Z)$ events, which could be correlated to high $p_T(H)$. Thus may cause deficit of events at low $p_T(\gamma\gamma)$. Apply similar cuts to madgraph sample and see if we observe shift in distribution.
- Use reco-level objects in the madgraph samples to perform analysis once set up the event selection.
- How to reconstruct ttH? May have to use variables which have only one reconstructed top quark e.g. m_{tH} or $p_T(tH)$. What differential C_1 values can be reached?

1.4 Other stuff

- Set up github repository to store plots related to trilinear analysis and also an ReadMe or useful code that I have written.
- Effect of trilinear reweighting in decay? Will this contribute to $p_T(\gamma\gamma)$ spectrum, thus effecting κ_λ value extracted from fit? Mentioned in [2] as δBR, can we just include as an input in combine tool? Worth asking authors about effect and if we can use value in paper.

2 References

[1] Fabio Maltoni, Davide Pagani, Ambresh Shivaji, and Xiaoran Zhao. Trilinear Higgs coupling determination via single-Higgs differential measurements at the LHC. *Eur. Phys. J.*, C77(12):887, 2017.

[2] Giuseppe Degrassi, Pier Paolo Giardino, Fabio Maltoni, and Davide Pagani. Probing the Higgs self coupling via single Higgs production at the LHC. *JHEP*, 12:080, 2016.