



CEPC NOTE

October 27, 2017



Measurement of $BR(H \rightarrow \gamma\gamma)$ via $Z \rightarrow q\bar{q}$ channel in e^+e^- collider at 250 GeV

The CEPC Collaboration

Abstract

We examine the prospects of measuring $BR(H \rightarrow \gamma\gamma)$ with a 125 GeV Higgs boson by measuring the cross section of the process $e^+e^- \rightarrow ZH \rightarrow q\bar{q}\gamma\gamma$ at the future Circular Electron Positron Collider (CEPC), assuming a integrated luminosity of 5ab^{-1} at center-of-mass energy of 250 GeV. Full simulations with the current CEPC model (CEPCv1) are used to determine the efficiency and shape of the signal and background data, and Monte Carlo (MC) simulations are used to determine the uncertainty of the measurement. A relative uncertainty of 15.6% can be achieved for the measurement of $\sigma(ZH) \cdot BR(H \rightarrow \gamma\gamma) \cdot BR(Z \rightarrow q\bar{q})$. The uncertainty's dependence on the photon reconstruction efficiency and calorimeter performance is also studied.

E-mail address: li.gang@mail.ihep.ac.cn

© Copyright 2017 IHEP for the benefit of the CEPC Collaboration.

Reproduction of this article or parts of it is allowed as specified in the CC-BY-3.0 license.

1 Introduction

In 2012, the ATLAS and CMS Collaborations discovered a 125 GeV boson consistent with the standard model Higgs boson [1, 2]. To fully understand the fundamental couplings associated with the Higgs boson, more precision measurement has to be made in future e^+e^- colliders [3]. As the Higgs boson is regarded as a portal to new physics beyond standard model, it is important to carefully compare the Higgs couplings to standard model predictions. The measurement of the Higgs di-photon branching ratio $BR(H \rightarrow \gamma\gamma)$ is an interesting candidate for this purpose, since not only does the process have a clean final state topology, but it may also contain loop contributions from new charged particles. However, the fact that this process is mediated by a particle loop causes the branching ratio to be relatively small. At Higgs mass $M_H = 125$ GeV, it reaches the maximum of about 0.23% [4].

In this paper we study an important channel involved in the total measurement of $BR(H \rightarrow \gamma\gamma)$, the Z boson di-jet decay channel in the Higgsstrahlung process, namely, $e^+e^- \rightarrow ZH \rightarrow q\bar{q}\gamma\gamma$. Analysis is performed based on full simulation using the detector model CEPCv1 described in section 3, making this study compliment to [5]. We assume a 125 GeV Higgs boson, with 250 GeV center-of-mass energy collision of integrated luminosity of 5ab^{-1} . The statistical precision of the measurement is determined by first casting the fully-simulated and reconstructed signal and background event in the (Higgs-decayed) di-photon invariant mass space and then performing Monte Carlo simulations based on the distribution to obtain the uncertainty of the measurement. We may also obtain uncertainty of the uncertainty reported by the fit program in signal fitting with multiple runs. Alternatively, one can use event counting $S/\sqrt{S+B}$ as a gauge for precision, where S and B , respectively, are the number of signal and background events in a small region around the signal peak ($\sim 1\sigma$) in the di-photon invariant mass space. The event counting measure is a good indicator of the final uncertainty (from MC simulations), and is used to optimize event selection scheme since it is much faster to evaluate. The methods used in this study can be easily generalized to the measurement of $BR(H \rightarrow \gamma\gamma)$ via other channels such as $e^+e^- \rightarrow ZH \rightarrow \mu\bar{\mu}\gamma\gamma$, as the analysis relies heavily on the Higgs-decayed di-photon information.

The paper is organized as follows: In section 2 we discuss the signal and relevant background processes and their theoretical detector responses. In section 3 we describe the CEPCv1 detector model used in simulation (and the reconstruction algorithm). In section 4 we present the event selection scheme and the determination of the uncertainty of the $\sigma(ZH) \cdot BR(H \rightarrow \gamma\gamma) \cdot BR(Z \rightarrow q\bar{q})$ measurement. In section 5 we further study the relation between the photon reconstruction algorithm and the measurement uncertainty.

2 Signal and background

In e^+e^- collision, the Higgs boson can be produced via Higgsstrahlung $e^+e^- \rightarrow ZH$, and the W or Z boson fusion $e^+e^- \rightarrow \nu_e\bar{\nu}_e H$ and $e^+e^- \rightarrow e^+e^- H$ [6, 7, 8]. At $\sqrt{s} = 250$ GeV and $M_H = 125$ GeV, the Higgsstrahlung process has a cross section of approximately 200 fb, much larger than the W and Z fusion production of the Higgs. Moreover, the hadronic decay of the Z boson also has a high branching ratio of about 70% in this energy level, thus the process $e^+e^- \rightarrow ZH \rightarrow q\bar{q}\gamma\gamma$ consists a major part of the processes involved in measuring $BR(H \rightarrow \gamma\gamma)$. Assuming $BR(H \rightarrow \gamma\gamma) = 0.23\%$, and an integrated luminosity of 5000fb^{-1} in 10 years of data taking in CEPC, we expect in total of approximately 1633 signal events.

The general tree-level background for $e^+e^- \rightarrow q\bar{q}\gamma\gamma$ consists of quark pair productions via photon and Z boson, with additional photons coming from initial state radiations (ISR) and final state radiations (FSR). The relevant diagrams agrees with the diagrams listed in [9]. The reducible background processes in which the quarks are produced via photon will be essentially eliminated using invariant mass information and our event selection scheme, and so are processes with no ISR since the Z boson would

be significantly off-shell. Otherwise, two energetic ISR photons or one ISR and one FSR photon may mimic the Higgs-decayed photons, while the quark pair also comes from an on-shell Z boson. We shall therefore ignore the reducible background events and focus on irreducible background where the photons mimics decay products of a particle recoiled by the on-shell Z boson.

The Monte Carlo data used for signal and background events are generated by the Whizard 1.95 package [10]. Due to the enormous amount of background events with low energy photons that can be easily eliminated in event selection, a pre-selection is performed on the MC sample before being taken to full simulation:

- The two most energetic photons with $\cos \theta_\gamma < 0.99$ must both exceed 25 GeV in energy. (θ_γ is the polar angle.)
- The two photons must satisfy $(\cos \theta_{\gamma_1} + 1)(\cos \theta_{\gamma_2} - 1) < -0.05$ and $(\cos \theta_{\gamma_1} - 1)(\cos \theta_{\gamma_2} + 1) < -0.05$
- The invariant mass of the two photons must satisfy $100 \text{ GeV} < M_{\gamma\gamma} < 160 \text{ GeV}$

The pre-selection imposed here are roughly on the same variables as the event selection imposed later to avoid unexpected deficiencies. To make sure that no events excluded in the pre-selection may enter the final cut in a full simulation due to calorimeter uncertainties, the final event selection boundaries must be more than 3σ away from the pre-selection bounds, where σ is the uncertainty introduced in the full simulation (principally from calorimeter) on the interested variable. Specifically, the standard deviation of the momentum direction of reconstructed photons is less than 0.0004 rad so that the uncertainty of $(\cos \theta_{\gamma_1} \pm 1)(\cos \theta_{\gamma_2} \pm 1)$ is about $0.0008 \oplus 0.0008 \approx 0.0012$. The uncertainty of $M_{\gamma\gamma}$ is at most (for 160 GeV) $2.0 \oplus 2.0 \approx 2.9 \text{ GeV}$ for Ecal performance of relative uncertainty $16\% / \sqrt{E}$. The first pre-selection is just to ensure that selected area has 100% Ecal coverage. To reflect the step-wise efficiency of the event selection, another "fast simulated" background sample is produced by imposing only the first of the three pre-selection criteria and then adding a smear to the photon energies to reflect the Ecal uncertainty, with the total efficiency is then adjusted to match the full simulated sample. This "fast simulation" sample is only used in evaluating the efficiency of the event selection.

3 Simulation and reconstruction

3.1 General simulation and reconstruction

[Generation: Whizard 1.95. Simulation: Mokka. Reconstruction: ArborV3KD. Detector Geometry: CEPCv1(?)] [The low-end resolution of Ecal is reflected by discarding all photons with energy less than 1 GeV. (Other low energy photons useful in photon recovery.)]

3.2 Photon recovery

Conversion of high energy photons that decayed from Higgs bosons into fermion pairs before reaching the detector is a serious reduction to the reconstructed event number. The general reconstruction algorithm, is not specifically designed for the $ZH \rightarrow q\bar{q}\gamma\gamma$ process that involves complex environment such as particle showers. It is necessary to develop a photon recovery algorithm specific for this study.

The general reconstruction yields a reconstruction rate of approximately 90% for photons decayed from Higgs boson in signal events. This means 80% of the events have both photons coming from the Higgs reconstructed, while another 19% of the events has one photon recovered. Since the majority of missing photon is due to conversion into e^+e^- pairs, we need to identify e^+e^- pairs that are close to each other (small track distance) that are consistent with a high energy photon. Partially motivated by the photon recovery algorithm used in [9], the following criteria for recovering a pair of electron is adopted:

For all pairs of particles with opposite charges:

- The angle between the 3 momentums of the particle pair should be less than 19° .
- The invariant mass of the particle pair should be less than 5 GeV.
- There should not be any other visible particles within the 7° cone of the converted photon's momentum (total of the particle pair).

(Selecting all charged particles as electron/positron candidate is due to the relatively large electron pion mistag rate in presence of jets.) The angle of between the 3-momentum is allowing since this criterion does not affect the recovery by much. The third criterion is for preventing false identification, especially from the particle shower. Although it may sacrifice some signal event whose Higgs-decayed photons happens to line up with some hadron debris, the false positive rate of the recovery algorithm is kept low by the final criterion.

After applying the recovery algorithm, the reconstruction rate of a single photon is boosted to $> 95\%$, with the success rate for reconstruction of both photons about 92.5% . We shall assume this recovery rate in the following analysis. The measurement uncertainty's dependence on the recovery rate is studied later in section 5.

4 Measurement of $\sigma(ZH) \cdot BR(H \rightarrow \gamma\gamma) \cdot BR(Z \rightarrow q\bar{q})$

4.1 Event selection

The measurement of $\sigma(ZH)BR(H \rightarrow \gamma\gamma)BR(Z \rightarrow q\bar{q})$ is based on the invariant mass information of the two energetic photons presumably decayed from the Higgs boson. We first select two photons with the highest invariant mass from a reconstructed event (after the additional photon recovery), and then apply a series of cut to filter out unwanted backgrounds. In limited Ecal precision, the di-photon momentum information proves to be more useful than the jet information from the quarks with more calorimeter hits, despite the latter's use in filtering out reducible background.

The variables being used in the event selection is selected by the following guideline. In $e^+e^- \rightarrow ZH \rightarrow q\bar{q}\gamma\gamma$ If we force the Higgs boson to be on-shell, the dip-photon system of the signal events has exactly one degree of freedom less than the irreducible background's because of this additional on shell requirement. However, it is important not to exploit this different in event selection since it would cause the background to mimic the signal in the di-photon invariance mass space. We should look for variables related to the external geometry (e.g. the z axis) where background events have different concentration from the signal events. The polar angles is a good choice since the ISR concentrate around the z axis. Taking into account that the two energetic ISRs are likely to be in opposite directions, a two dimension cut is necessary. We then apply cuts on the angle of the two photons with the recoil momentum to further filter background events based on the paraxial feature of the ISR. We finally apply a two dimensional cut on the energy-invariant mass plane of the di-photon system to single out the signal peak in the invariant mass space. After selecting the cut schemes, we then optimized the parameters using Monte Carlo simulated background and full-simulated signal, using the event counting criterion $S/\sqrt{S+B}$ near the signal peak ($125 \text{ GeV} \pm 5 \text{ GeV}$), which reflects the uncertainty of the measurement as mentioned before. The efficiency of the cut chain carries over to fully-simulated background events. The cut chain is summarized here:

- Two dimensional cuts on the polar angle: $(\cos \theta_{\gamma_1} + 1)(\cos \theta_{\gamma_2} - 1) < -0.07$ and $(\cos \theta_{\gamma_1} - 1)(\cos \theta_{\gamma_2} + 1) < -0.07$ and $-0.9 < (\cos \theta_{\gamma_1} + \cos \theta_{\gamma_2})$ and $(\cos \theta_{\gamma_1} + \cos \theta_{\gamma_2}) < 0.9$
- Cuts on angle with recoil momentum: $\cos(p_{\gamma_1}, p_{\text{recoil}}) > -0.95$ and $\cos(p_{\gamma_2}, p_{\text{recoil}}) < 0.70$

- Two dimensional cuts on the E-M plane of the di-photon system: $0.74 \cdot M_{\gamma\gamma} + 41 \text{ GeV} < E_{\gamma\gamma} < 0.48 \cdot M_{\gamma\gamma} + 83 \text{ GeV}$

Moreover, we focus on the window $113 \text{ GeV} < M_{\gamma\gamma} < 150 \text{ GeV}$ to perform the fitting. After applying the event selection criteria, around 828 (50.7%) signal events survived, while 29137 (0.265%) background survived in the fitting window. (Percentage calculated with expected event number before pre-selection). In a small regions around the signal (about 1 sigma), around 581 (59.13%) signal events survived while 6177 (3.16%) background events survived. The efficiency for signal events is evaluated using a sample much larger than the actual event counting 1633 (used: 40000), so the uncertainty of the efficiency is negligible compared to the actual uncertainty in fitting the signal peak. (The background efficiency is evaluated with a sample comparable to the actual size 11011914 (used: 11300000).) Detailed step-wise efficiency are presented in Table 1 for a small region around the signal and Table 2 for the total and fitting region efficiencies.

Table 1: Event selection efficiencies in the region $122 \text{ GeV} < M_{\gamma\gamma} < 128 \text{ GeV}$ (near signal).

After:	Uncut	Precut	Cut1	Cut2	Cut3
Signal:	982	976	754	654	581
Efficiency:	100%	99.30%	76.79%	66.56%	59.13%
Background:	195743	17272	15983	9644	6177
Efficiency:	100%	8.82%	8.17%	4.97%	3.16%
$S / \sqrt{S + B} = 7.07 \quad (S = 581 \quad B = 6177)$					

Table 2: Event selection efficiencies (total & in fitting window $113 \sim 150 \text{ GeV}$).

After:	Uncut	Precut	Cut1	Cut2	Cut3	Fitting window
Signal:	1633	1547	1161	1000	838	828
Efficiency:	100%	94.77%	75.03%	61.28%	51.33%	50.70%
Background:	11011914	159657	141860	83338	47022	29137
Efficiency:	100%	1.45%	1.29%	0.757%	0.427%	0.265%

The filtered events, casted in the di-photon invariant mass space is presented in Fig. 1. The signal and background distributions are independently fitted to as a truth reference of the Monte Carlos evaluation of the uncertainty.

4.2 Evaluation of uncertainty

In the actual experiment, we obtain a similar distribution after applying the cut chain and compare the distribution to this simulation based on the Standard Model. To obtain the signal event number, we apply a fit to the combined distribution of the signal and background events, and the uncertainty of the signal events comes from that of the fitting. Assuming we will obtain an (signal) distribution close to the Standard Model prediction, we can measure the uncertainty by attempting to perform a fitting on Monte Carlo generated distributions based on the given distribution. The background is relatively smooth, so a quadratic function is used as a fit function. The signal width comes mainly from the Ecal errors, and the actual Higgs width is obscured, so a simple Gaussian distribution is used. The fit to a distribution is done by the Minuit algorithm in the Root program (v6.04/16) [11, 12], for two times, while the initial values are set close to the truth value. The uncertainty stabilizes as 10000 Monte Carlo distributions are fitted. The measurement is plotted in Fig. 2 (a), and the uncertainty is 15.4%. The uncertainty reported by the program is plotted in Fig. 2, with a mean value of 15.9%, and an uncertainty of $\pm 0.06\%$. However, the

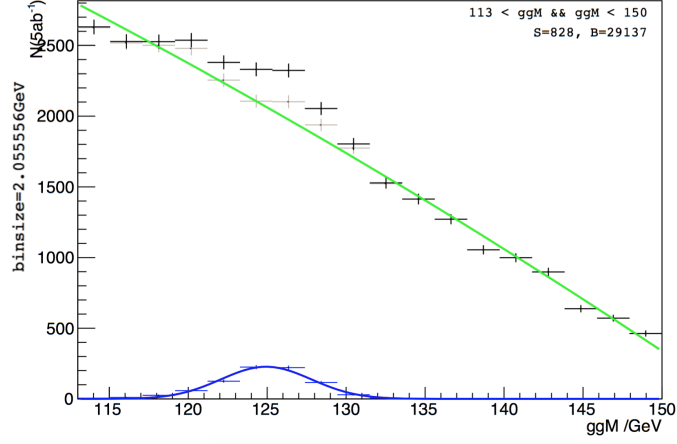


Figure 1: **Di-photon invariant mass distribution.** The blue line and green line are distributions fitted independently for signal and background distribution for MC evaluation of uncertainty.

mean value of the measurement systematically deviates from the truth value by about the 2.7%, whose direct sum with the statistical uncertainty $15.4\% \oplus 2.7\% \approx 15.6\%$ the reported uncertainty of 15.9%. Since the value of $\sigma(ZH)$ and $BR(Z \rightarrow q\bar{q})$ is known precisely from electroweak precision measurement, the measured uncertainty of $BR(H \rightarrow \gamma\gamma)$ is simply that of $\sigma(ZH) \cdot BR(H \rightarrow \gamma\gamma) \cdot BR(Z \rightarrow q\bar{q})$.

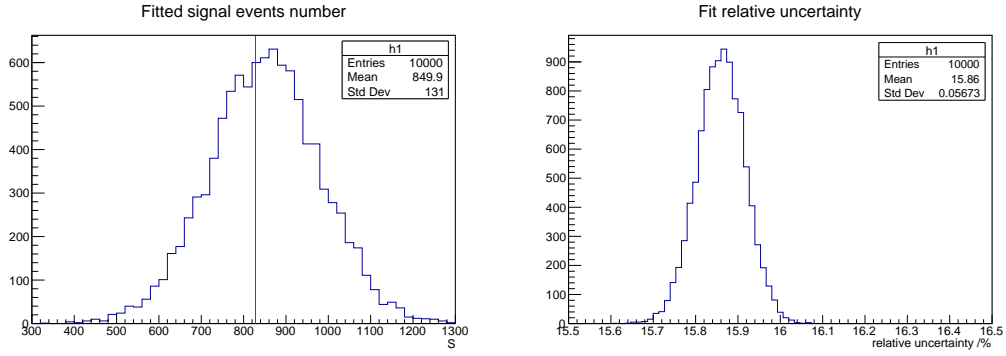


Figure 2: **Distribution of measurements and uncertainties in Monte Carlo simulation of the fitting.** (a) contains the expected signal count corresponding to the fit. The red line in (a) indicates the truth value. A uncertainty of 15.4% can be achieved. (b) contains the fitting uncertainties (of the signal count) reported by the fitting program. The mean value 15.6% is in agreement with the standard deviation of (a). The uncertainty of the measurement uncertainty is $\pm 0.06\%$. [Update!]

5 Relation to photo reconstruction algorithm

We also study the dependence of our measurement's uncertainty on the photon conversion recovery algorithm and variance in signal width due to calorimeter performance. The photon recovery algorithm used in this paper boosts the total reconstruction success rate from 81% to 92.5% (in addition to some rare cases where γ_1 and γ_2 are reversed, i.e. the energy comparison is reversed in the full simulation due to calorimeter uncertainty, which are deemed unsuccessful here). We explore the dependence of $\delta s/s$ ($s = \sigma(ZH) \cdot BR(H \rightarrow \gamma\gamma) \cdot BR(Z \rightarrow q\bar{q})$) on the efficiency of the photon recovery algorithm by simply

compensating back or further deducting certain amount of expected signal events in the Monte Carlo evaluation of the uncertainty. Dependence on the Ecal performance is also studied. The signal width (~ 2.9 GeV in previous analysis) that principally originate from Ecal is related to the direct sum of the uncertainty of the energy of the resonant photons, each of which is described by $\frac{r\%}{\sqrt{E_i}} \cdot E_i$, where r is the scaled relative uncertainty of the Ecal, and E_i the energy of the respective photons. The total uncertainty is event-dependent, so we will simplify the picture and present the dependence of measurement uncertainty on the signal width. Fig. 3 shows the dependence of the uncertainty on the total reconstruction efficiency and the width of the signal distribution. It can be seen that uncertainty mildly decrease as the reconstruction efficiency grows. For signal width of about 2.9 GeV (as in previous analysis), the uncertainty can reach below 15% with more photons successfully reconstructed. The dependence on the signal width is much larger, apparently. For the level of reconstruction efficiency in this analysis, the uncertainty ranges from about 25% to below 10% as the signal width varies from 5 GeV to 1 GeV.

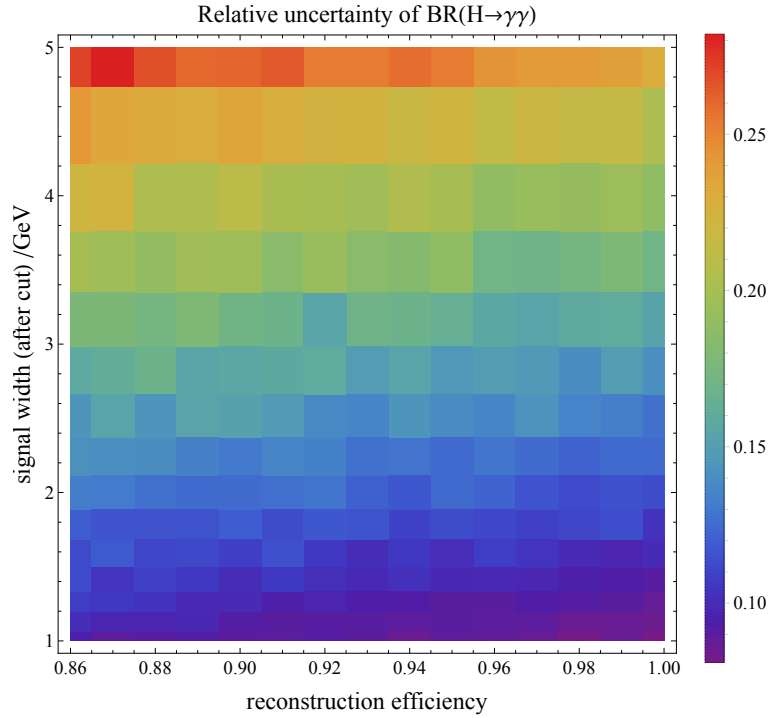


Figure 3: **Dependence of relative uncertainty on total photon reconstruction efficiency and width of signal distribution.** Color represents relative uncertainty of the measurement $\delta s/s$, where $s = \sigma(ZH) \cdot BR(H \rightarrow \gamma\gamma) \cdot BR(Z \rightarrow q\bar{q})$.

References

- [1] ATLAS Collaboration, G. Aad et al., Phys. Lett. **B716** (2012) 1–29, arXiv:1207.7214 [hep-ex].
- [2] CMS Collaboration, S. Chatrchyan et al., Phys. Lett. **B716** (2012) 30–61, arXiv:1207.7235 [hep-ex].
- [3] ECFA/DESY LC Physics Working Group Collaboration, E. Accomando et al., Phys. Rept. **299** (1998) 1–78, arXiv:hep-ph/9705442 [hep-ph].

- [4] S. Actis, G. Passarino, C. Sturm, and S. Uccirati, Nucl. Phys. **B811** (2009) 182–273, arXiv:0809.3667 [hep-ph].
- [5] F. Wang, G. Li, M. Ruan, X. Zhou, and Z. Zhang, ??? (201?).
- [6] S. L. Glashow, Nucl. Phys. **22** (1961) 579–588.
- [7] S. Weinberg, Phys. Rev. Lett. **19** (1967) 1264–1266.
- [8] A. Salam and N. Svartholm, Ed. N. Svartholm, Stockholm, \ Almquist and Wiksell **367** (1968).
- [9] E. Boos, J. C. Brient, D. W. Reid, H. J. Schreiber, and R. Shanidze, Eur. Phys. J. **C19** (2001) 455–461, arXiv:hep-ph/0011366 [hep-ph]. [,1234(2000)].
- [10] W. Kilian, T. Ohl, and J. Reuter, Eur. Phys. J. **C71** (2011) 1742, arXiv:0708.4233 [hep-ph].
- [11] R. Brun and F. Rademakers, Nucl. Instrum. Meth. **A389** (1997) 81–86.
- [12] F. James and M. Roos, Comput. Phys. Commun. **10** (1975) 343–367.

Appendices

A Step-wise di-photon invariant mass distribution

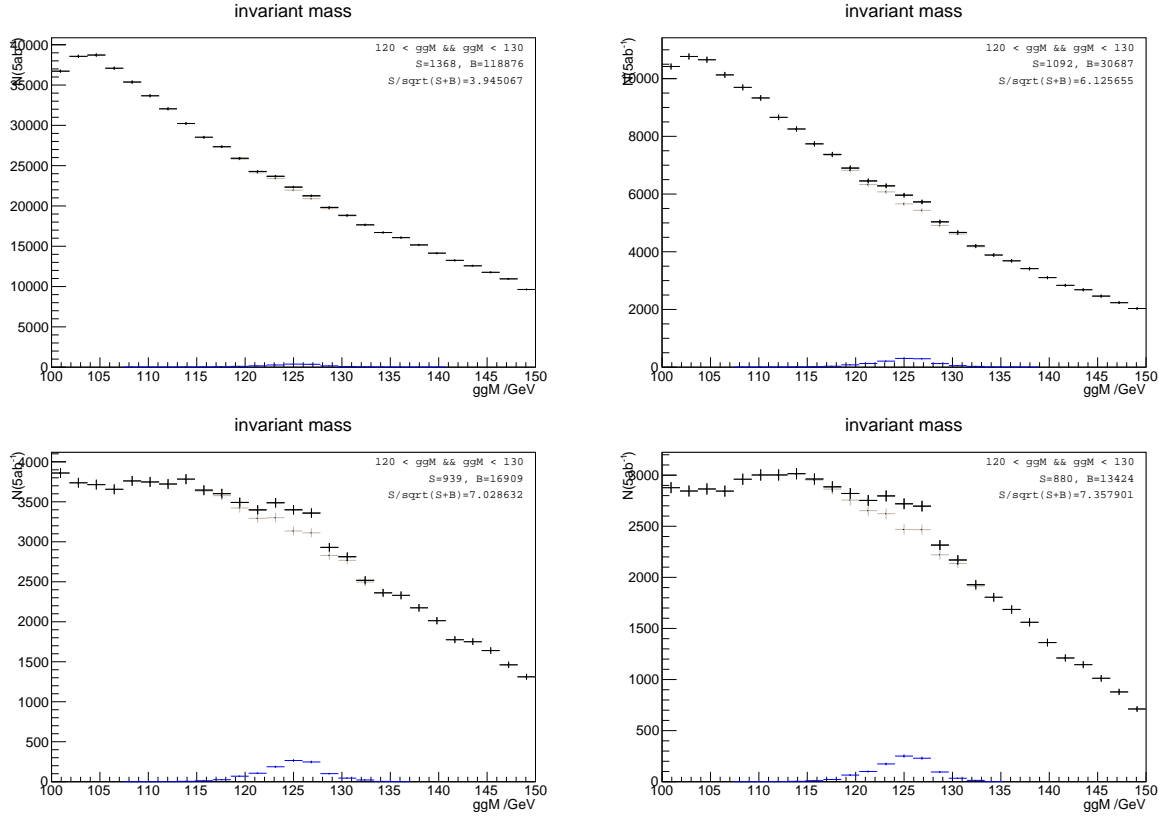


Figure 4: **Signal and background distribution in the di-photon invariant mass dimension:** after 1) Precut, 2) Cut1, 3) Cut2, 4) Cut3.