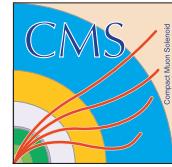




# Trigger Strategies for $H \rightarrow \gamma\gamma$ Channel in the CMS Detector at the $\sqrt{s} = 13$ TeV LHC



---

*Joshua Hardenbrook - Princeton University*

*October 22<sup>nd</sup>, 2013*

*In Fullfillment of the Princeton Physics PhD Experimental  
Project Requirement*

*Advisor: Professor Christopher Tully*

---

# Table of Contents

---

## CONTENTS

1.	Introduction	3
1.1.	The Large Hadron Collider	4
1.2.	The CMS Detector	4
1.3.	The CMS Trigger System	4
2.	The Higgs Signal and Backgrounds	6
3.	The High Level Trigger (HLT) Definitions	8
3.1.	Coordinate Conventions	8
3.2.	Photon Kinematic Variables	9
3.3.	Photon Isolation Variables	11
3.4.	Photon Showershape Variables	12
3.5.	Summary Photon Identification	13
4.	HLT definitions	14
5.	Analysis Considerations	14
5.1.	OpenHLT Simulation and Rate Analysis	15
5.2.	Tag and Probe	15
5.3.	Sample Run 208390	15
6.	Conclusions	18
	References	18

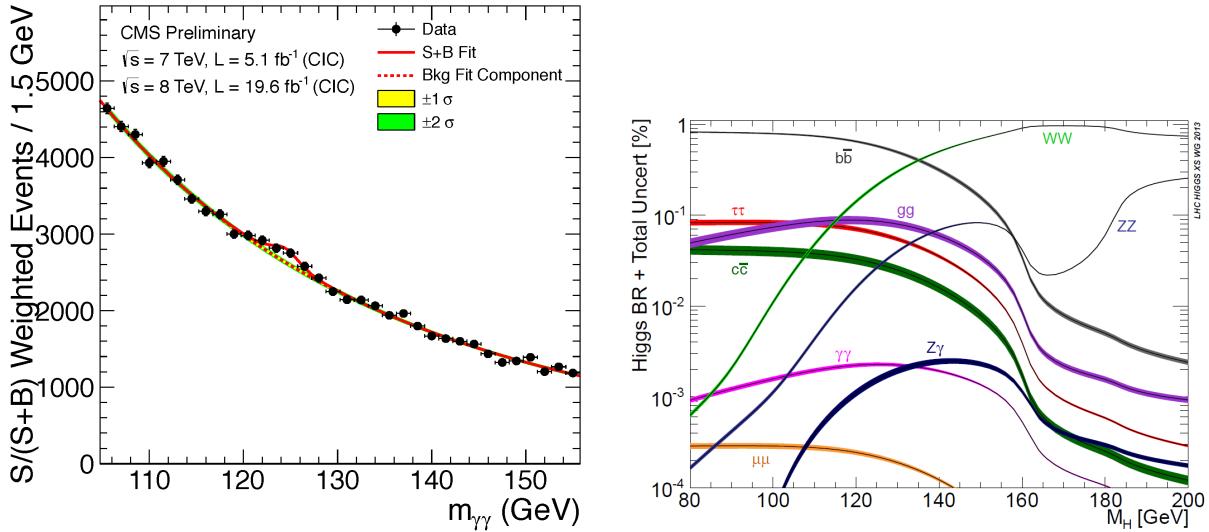


FIGURE 1. (left) The signal weighted invariant mass plot of the  $H \rightarrow \gamma\gamma$  analysis displaying an excess at  $m_{\gamma\gamma} \approx 125$  GeV over polynomial fit of the background distribution (right) The Higgs branching ratios as a function of  $m_H$ .

## 1. INTRODUCTION

The Large Hadron Collider (LHC) is a proton-proton collider installed in the former LEP tunnel at CERN Laboratories on the border of Geneva, Switzerland and France. The LHC was built first and foremost as a discovery machine looking for new physics at the TeV scale.

On July 4, 2011 the Higgs Boson was declared as discovered by both the CMS and ATLAS collaborations. On October 8, 2013 the Nobel Prize in physics was awarded for the theoretical work of developing the Higgs. This award not only recognized the great insight of Peter Higgs and Francois Englert, but also the grand scientific and technological challenge met by the two experiments and LHC coordination.

The production of the Higgs resonance is rare and its decay products difficult to extract from more prevalent background processes. The success of the numerous analyses leveraged the unique signature of the Higgs mass resonance differentiate it from its mimicing processes.

As the Higgs mass resonance width is small  $O(\text{MeV})$ , the precision mass resolution was a powerful tool. Accordingly, the highest contributing significance to the CMS discovery comes from the  $H \rightarrow \gamma\gamma$  channel where the CMS (Compact Muon Solenoid) Detector has high mass resolution from sub-precent calibration of the electromagnetic calorimeter to electromagnetic energy. The  $H \rightarrow \gamma\gamma$  process occurs 3 orders of magnitude less often than  $H \rightarrow b\bar{b}$  (the dominant decay mode). As these events held such great statistical value (after optimized vertex selection of the diphoton pair), it was imperative that every  $H \rightarrow \gamma\gamma$  produced was recognized and recorded by the machine. To implement this filtering process, the CMS detector uses a trigger system designed to quickly check if an event is interesting and if so, record it.

When the LHC begins colliding protons in 2015 at  $\sqrt{s} = 13$  TeV the event triggers for the diphoton analysis must be modified to account for the evolving goals and needs of the  $H \rightarrow \gamma\gamma$  analysis as well as the entirety of the CMS physics program. We must account and prepare for a number of uncertainties in future collisions. The pileup from multiple interaction vertices will not be known, the luminosity will be at least a factor 2 higher, and the increase in  $\sqrt{s}$  will all present necessary changes to the  $H \rightarrow \gamma\gamma$  analysis.

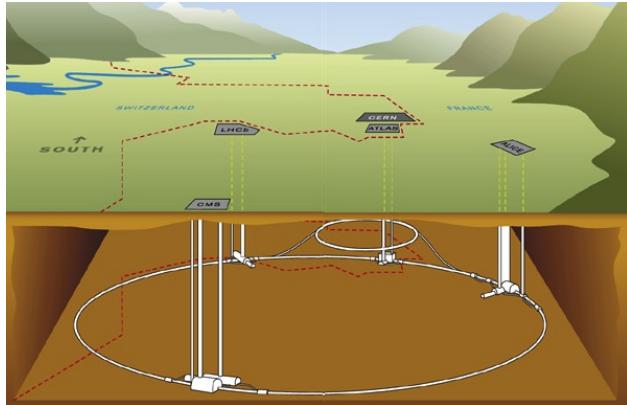


FIGURE 2. The LHC tunnel installed on the border of Geneva, Switzerland and France. The experiments are distributed across the circumference of the ring.

Discovery is no longer the driving force of the  $H \rightarrow \gamma\gamma$  analysis and priorities will be shifted towards new goals: measurement of the signal strength,  $\sigma \times$  branching ratio, spin studies, higgs tagging, self-couplings, etc.

In this project, we present future trigger strategies for the 13 TeV LHC based on modifications to the latest  $\sqrt{s} = 8$  TeV trigger suite for the  $H \rightarrow \gamma\gamma$  analysis. We consider modifications to the current analysis workflow to make significant improvements in trigger rate without sacrificing signal sensitivity. Furthermore, we make broad scans of the HLT rate space in preparation for luminosity related threshold changes that must be iteratively made to the triggers.

**1.1. The Large Hadron Collider.** The bunches of protons in the LHC are bent into a circular trajectory by more than 1200 superconducting dipole magnets and are focused and maintained close to the ideal orbit around the ring by hundreds of superconducting quadrupole magnets. Thousands of corrector magnets around the ring allow the beam to be steered closer to the ideal orbit, make the focusing independent of the particles energy variations within a bunch, and cancel the effects of higher order multipoles in the fields induced by small field imperfections in the main magnets. The radiofrequency (RF) field in superconducting cavities placed periodically around the ring accelerates the protons from the injection energy of 450 GeV to the final operating energy, which is designed to be 7 TeV per beam. The RF field also causes the protons to be bunched, as only particles at or near a certain equilibrium phase on the RF wave will be accelerated stably. Special quadrupoles around each interaction region focus the bunches down to a small transverse size, to increase the likelihood of a proton-proton collision each time two bunches pass through each other.

**1.2. The CMS Detector.** The Compact Muon Solenoid (CMS) Detector is a general-purpose detector consisting of an all silicon tracker, a precision electromagnetic calorimeter (ECAL), a hadron calorimeter (HCAL), a 4 T superconducting solenoid and muon chambers. The solenoid deflects charged particles whose paths are traced in the tracker, making it possible to reconstruct the particles momentum. The two calorimeters reconstruct the energy of and identify photons, electrons and hadronic jets. As shown in Figure 3 the detector has cylindrical symmetry about the interaction point where the proton beams collide. By maintaining near full coverage of the interaction point it is possible to detect signatures such as neutrinos or other weakly interacting particles as missing energy.

### 1.3. The CMS Trigger System.

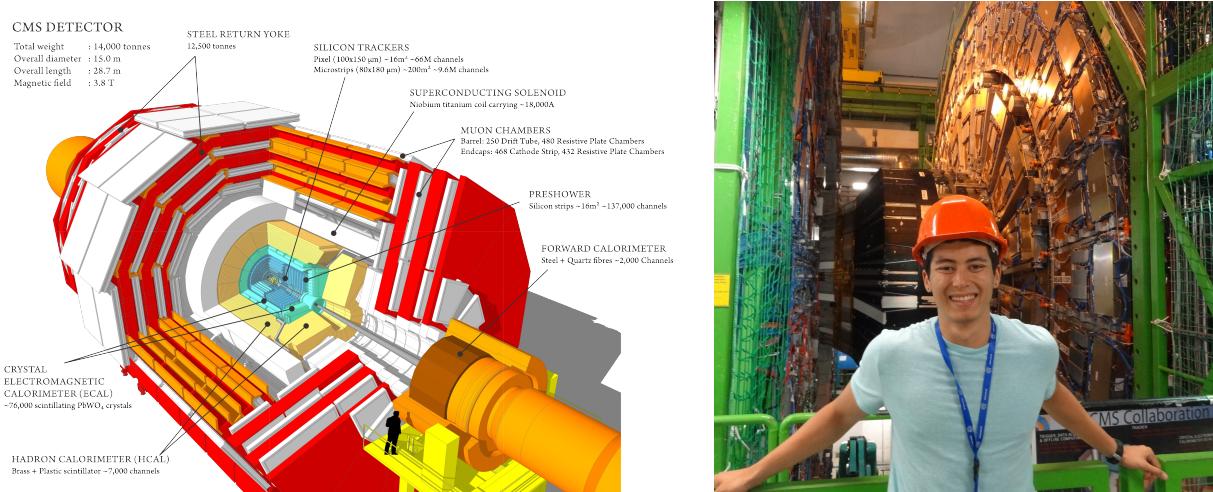
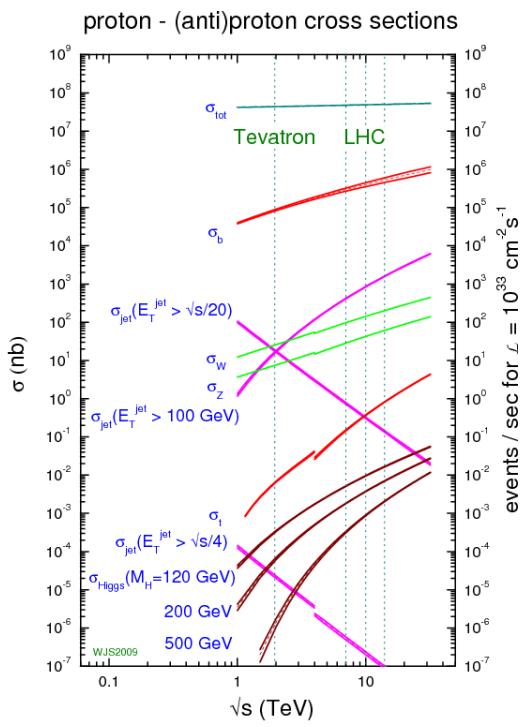


FIGURE 3. (left) A tear-away view of the inner detectors of CMS (right) The CMS endcap currently detached from the inner barrel for upgrades during the shutdown



Common cross sections of proton collisions as a function of the center of mass energy  $\sqrt{s}$

The CMS Trigger System exists as a filter through which events are determined to be “interesting”. It is both unnecessary and inefficient to record anything that occurs in the detector electronics. Most events that occur from colliding protons are well understood. To the left, you can see the logarithmic plot of common physics processes for proton-proton scattering. Events such as the production of a  $b$  quark occur at  $\approx 10^6$  Hz at a luminosity of  $\mathcal{L} = 10^{33} \text{ cm}^{-2}$  whereas the production of the Higgs is much lower at  $\approx 10^{-2}$  Hz.

At design luminosity, the LHC has beam crossings at a rate of  $\approx 40$  MHz with each crossing coming spaced at  $\approx 25$  ns. For each crossing there are  $\approx 20$  inelastic collisions (referred to as pile up) contained in an event file of  $\approx 1$  Mb. However the bandwidth for storage is limited to  $\approx 10^2$  Hz and equivalently  $10^2$  Mb/s. Generally all but one of the inelastic collisions is interesting and a large excess of uninteresting activity is generated in the detector electronics. The trigger must be robust enough to select this needle in a haystack event while remaining computationally efficient in maximizing the limited bandwidth.

The CMS Trigger system is designed to read events at the event crossing frequency and generate the factor  $10^5$  of rejection between the crossing frequency and the archival capacity. This factor is far too large to achieve in a single step given the complexity of triggers and event reconstruction. Therefore split into two steps: The Level 1 (L1) and High Level (HLT) Trigger systems.

The  $O(10^7)$  events per second first pass through the L1 Trigger which reads out events at  $10^5$  Hz. From here, the High Level Trigger makes the final decision as to which events are kept. Approximately 350 Hz is

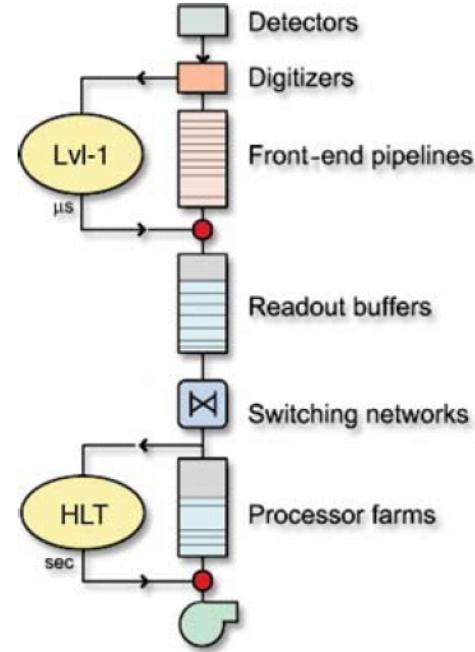
processed and stored, 300 Hz is “parked” (stored but processed later), and 1kHZ is partially stored (only the HLT level information and not the RAW detector information) and used for data scouting for future analysis.

The most basic criterion for interesting events are hard physics events with high momentum transfer,  $Q^2$ . As the protons collide with effectively no transverse momentum, any event with significant deposits of transverse energy (or even missing transverse momentum) is indicative of a hard physics process. The number of objects with a given transverse momentum falls off exponentially, so a simple minded way to reduce the rate of processed events is to raise the threshold of accepted events.

More specific criterion for “interesting events” is analysis dependent. Generally, analyses are categorized by their final state signature. Thus, the trigger requires loose identification on the objects of that signature such as the isolation and shape of energy deposition. Once the event has passed the Level 1 and HLT Triggers, tighter and more computationally costly selection can be made unrestricted by archival limitations.

As there is a limited amount of bandwidth for processing the events, the numerous analyses of CMS are given a budget (measured in Hz) for the triggers they request. As it stands the Higgs Diphoton analysis is assigned a budget of 30 Hz for its diphoton trigger suite. As the diphoton channel was of high priority in the 7 and 8 TeV running this accounted for a significant fraction ( $\approx 10\%$ ) of the overall budget.

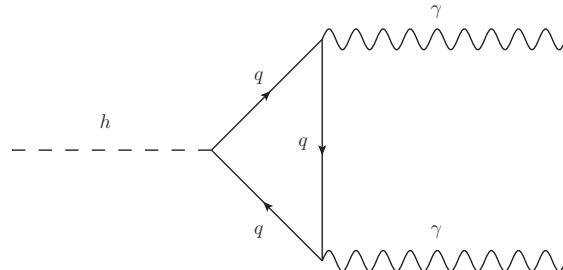
As the luminosity of the machine increases, we expect proportionally more events per second and must accordingly alter the triggers.



A diagrammatic representation of the level 1 and HLT trigger processing

## 2. THE HIGGS SIGNAL AND BACKGROUNDS

If we are going to understand how to trigger on  $H \rightarrow \gamma\gamma$  events we need to have a strong understanding of what we are looking for (signal) and what can fake what we are looking for (background).



The Higgs coupling to photons through a quark loop

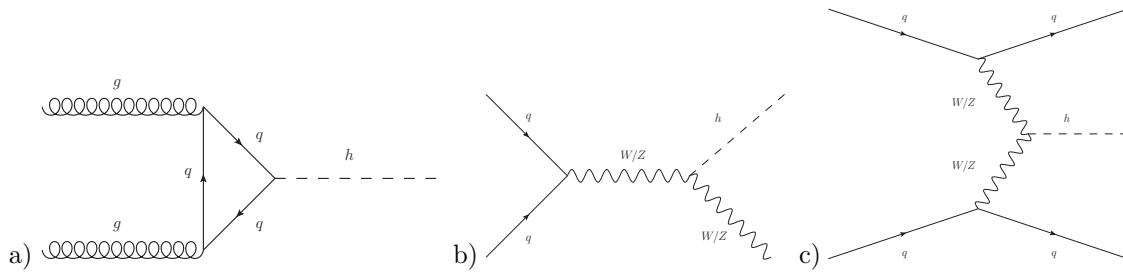


FIGURE 4. Example Feynman diagrams contributing to Higgs production. a) gluon fusion  
b) associated production c) vector boson fusion

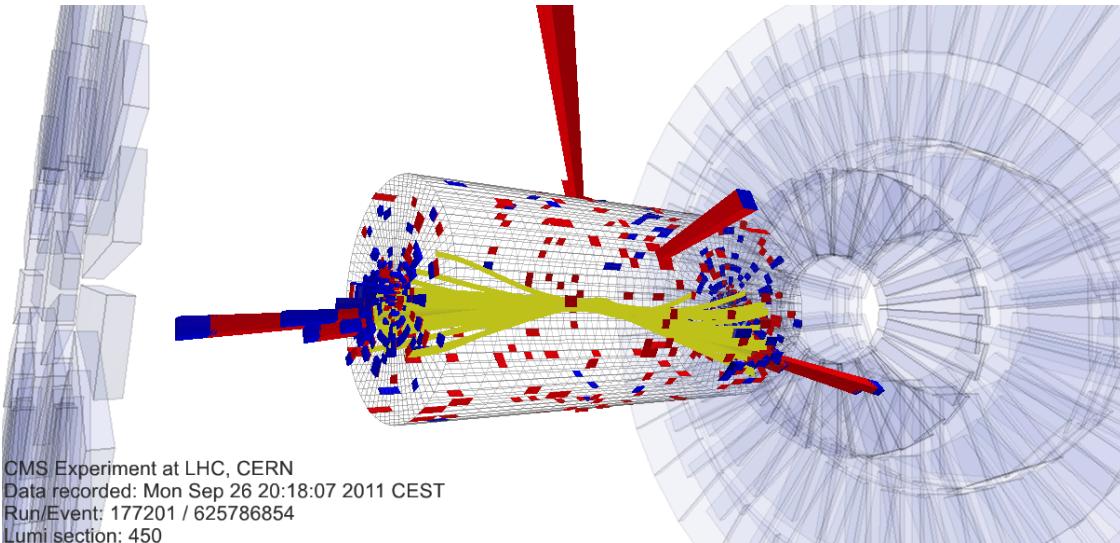


FIGURE 5. An unambiguous Higgs vector boson fusion (VBF) event recorded by the CMS experiment in September, 2012. Two central high  $p_t$  photons shown as two large red ecal deposits and two forward jets with their associated tracks highlighted in yellow

The Higgs diphoton signal is the combination of 3 dominant production modes: vector boson fusion (VBF), associated production (VH), and gluon fusion ( $ggH$ ) as shown in Figure 4. The Higgs then decays to diphotons through a heavy quark loop (generally a top) to diphotons, shown above. This additional loop suppresses the branching fraction to diphotons. This loop is necessary as the Higgs does not couple directly to photons, otherwise they would carry a mass.

Production Mode $\sqrt{s} = 8$ TeV	Cross Section $\times$ Branching Ratio $Br(H \rightarrow \gamma\gamma) = 2.29 \times 10^{-3}$
Vector boson fusion $qq \rightarrow qqH$	$\sim 3.5$ fb
Associated production $qq \rightarrow (W/Z)H$	$\sim 2.4$ fb
Gluon fusion: $gg \rightarrow H$	$\sim 44$ fb

The backgrounds to the  $H \rightarrow \gamma\gamma$  study fall into two categories: reducible, and irreducible. Reducible backgrounds are processes which look like photons, but are actually jets or leptons. Reducible backgrounds, as their name implies, can be effectively reduced by applying photon identification criteria. Irreducible backgrounds are processes that have true photons, but do not come from a higgs decay. These processes are more difficult to eliminate as we lose the handle of looking for qualities intrinsic to photons and must rely on kinematic relationships such as their invariant mass and the differential distribution of transverse momentums.

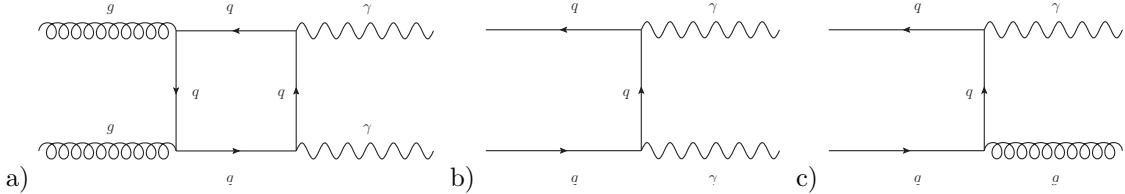


FIGURE 6. Example Feynman diagrams contributing to background. This is not an exhaustive list. a) Irreducible Box b) Irreducible Born c) Bremmstrahlung  $\gamma + \text{jet}$

Background Process (jets are EM enriched)	Cross Section $\sqrt{s} = 8 \text{ TeV}$
Reducible QCD multi-jets	$\sim 10^{10} \text{ fb}$
Reducible Bremmstrahlung $\gamma + j$	$\sim 10^7 \text{ fb}$
Irreducible Box: $gg \rightarrow \gamma\gamma$	$\sim 10^5 \text{ fb}$
Irreducible Born: $qq \rightarrow \gamma\gamma$	$\sim 10^5 \text{ fb}$

Despite the difficulties of irreducible backgrounds, their cross sections make their consideration second order to the overwhelming production of QCD multijets and Bremsstrahlung  $\gamma + \text{jet}$  (Figure ??).

Although the true signature of a jet and photon are vastly different in character, the overwhelming cross sections of QCD and  $\gamma + j$  produce highly electromagnetically enriched jets with character nearly indistinguishable from photons. A variety of derived quantities based on calorimeter and tracking information efficiently reject these signatures.

### 3. THE HIGH LEVEL TRIGGER (HLT) DEFINITIONS

To make sense the trigger bit names it is necessary to first define the variables in which it is parameterized. In the following subsections I will explain the coordinate space, kinematic variables, and derived photon identification criteria.

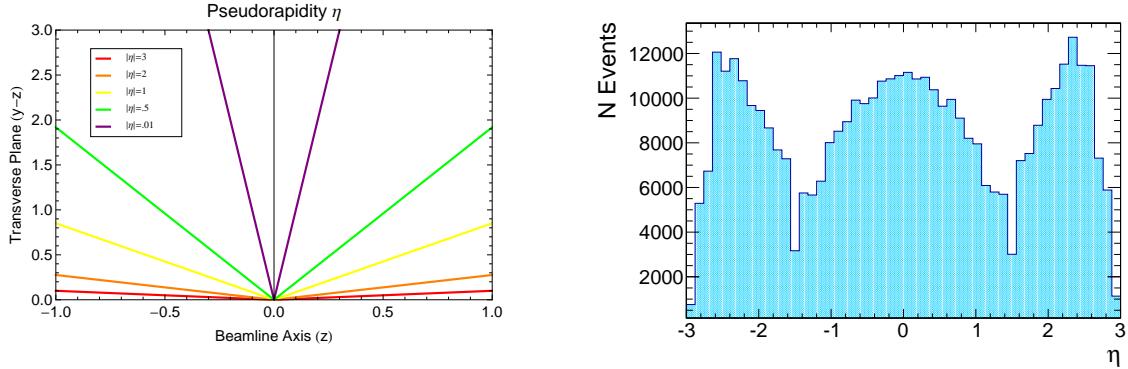
For each quantity, where applicable, I show the distribution of the HLT Trigger level information reconstructed in CMS Run 208390 contained in the photon primary dataset. The photon primary dataset consists of all events collected by photon related triggers (a superset of the higgs diphoton triggers). This run is the data for which our final result is expressed in the following section. As each trigger requires two photons, each variable will in general have two distributions. For reasons we will express after the HLT definitions, we refer to these as the seeded and unseeded “legs” or “photons” of the trigger.

#### 3.1. Coordinate Conventions.

**Pseudorapidity  $\eta$**  As the detector has cylindrical symmetry, the coordinate system used most commonly is two dimensional  $(\eta, \phi)$ . The pseudo-rapidity,  $\eta$  is an approximation to rapidity,  $y$ , that is exact in the  $\beta = 1$  limit:

$$y = \frac{1}{2} \log \frac{E - p_z}{E + p_z} \quad \eta = -\log \left( \tan \left( \frac{\theta}{2} \right) \right)$$

where  $\theta$  is the angle from the positive beam axis. This variable is useful for a number of reasons. Firstly, rapidity are invariant under longitudinal lorentz boosts along the beam axis. Also, for the energies being probed the particles in the decay products are of negligible mass and the approximation  $\eta \approx y$  is nearly exact. Given this relation, pseudorapidity provides an intuitive geometric interpretation. Near full solid angle coverage is provided within  $|\eta| < 5$

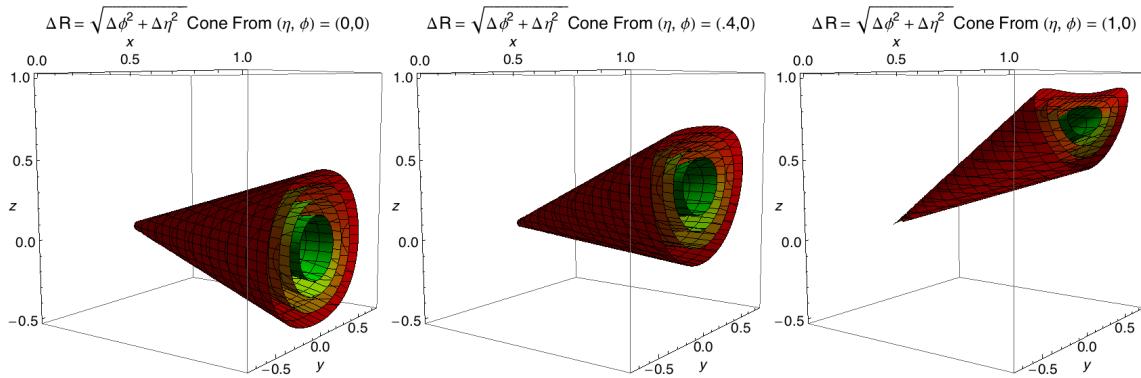


ΔR Given our coordinate system,  $\Delta R$  is our longitudinally boost invariant notion of distance:

$$\Delta R = \sqrt{(\Delta\phi)^2 + (\Delta\eta)^2}$$

Fixed values of  $\Delta R$  form a solid angle “cone” extending from the interaction point outward. This can be seen by using our definition of  $\eta$  to convert from cylindrical coordinates to  $(x, y, z)$  and consider the distance relative to the point  $(\eta_0, \phi_0)$

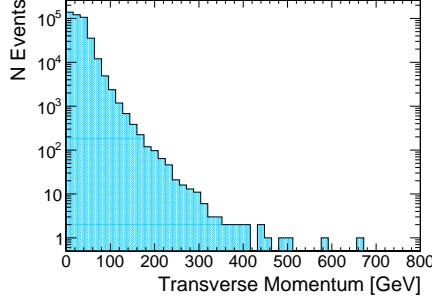
$$\Delta R = \sqrt{(\phi_0 - \tan^{-1}(y/x))^2 + \left(\eta_0 + \log\left(\tan\frac{\cos^{-1}(z/\sqrt{x^2+y^2})}{2}\right)\right)^2}$$



Contours of this equation are plotted above for  $\Delta R = 0.5, 0.4, 0.3, 0.2$  going from red  $\rightarrow$  green.  $\Delta R = .04$  is the most common size for generic, jet clustering. this metric is also invariant under longitudinal boosts because this quantity is defined in terms of a difference in pseudorapidity and  $\phi$ .

### 3.2. Photon Kinematic Variables.

Photon Transverse Energy  $p_T = \sqrt{p_x^2 + p_y^2}$ . The two partons engaged in the hard physics process have asymptotically zero transverse momentum before colliding. Thus, objects which have significant momentum contributions in the transverse plane are likely due to high  $q^2$  (momentum transfer) interactions. Conversely, low momentum transverse interactions, like scattering, will produce events in the forward (high  $|\eta|$ ) directions.



Invariant mass of the diphoton pair  $m_{\gamma\gamma}$  Naturally, now that the Higgs has been discovered, cutting on the mass is an effective way to select signal while suppressing background. Expressing the invariant mass in terms of the angle between the two photons and their energies:

$$m_{\gamma\gamma}^2 = 2E_1E_2(1 - \cos\theta_{12})$$

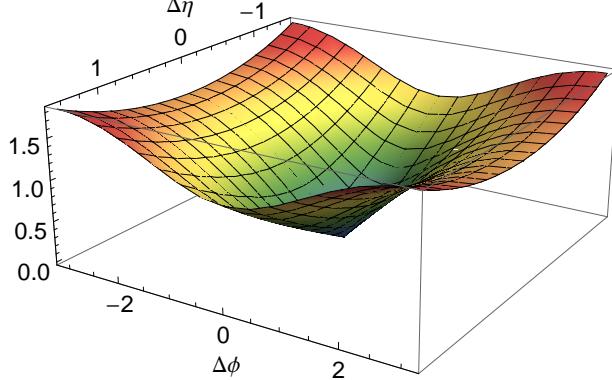
We can express the four-momentum in a convenient basis for the detector geometry as:  $p^\mu = (E, p_x, p_y, p_z) = (m_T \cosh y, p_T \sin \phi, p_T \cos \phi, m_T \sinh y)$  where  $m_T = \sqrt{m^2 + p_x^2 + p_y^2}$  and  $p_T = \sqrt{p_x^2 + p_y^2}$ . So in the massless limit we have :

$$p^\mu = p_T(\cosh \eta, \sin \phi, \cos \phi, \sinh \eta)$$

exactly since  $\eta = y$ . Using this form of the four-momentum, we can express the diphoton invariant mass spectrum in terms of the differences in  $\eta$  and  $\phi$ :

$$m_{\gamma\gamma}^2 = 2p_1^\mu p_2^\mu = 2p_t^1 p_t^2 (\cosh \Delta\eta - \cos \Delta\phi)$$

Diphoton Mass in Barrel for ( $p_{t_1} = p_{t_2} = 1$ )



which (as expected) is invariant under longitudinal boosts because it is written in terms of differences of rapidity and  $\phi$ . As we are searching for a heavy resonance we expect the majority of events to occur near  $\theta = 0 \implies |\eta| \approx 0 \implies \Delta\eta \approx 0$  maximizing the  $p_T$  of each photon with  $\Delta\phi \approx \pi$  i.e. back to back photons. Using these predictions as a rule of thumb, we can reduce the statement to:

$$m_{\gamma\gamma}^2 = 4p_T^1 p_T^2$$

Which we can re-express in terms of the scalar sum and difference of the  $p_T$ 's  $\Delta = (p_T^1 - p_T^2)/2$  and  $\Sigma = (p_T^1 + p_T^2)/2$

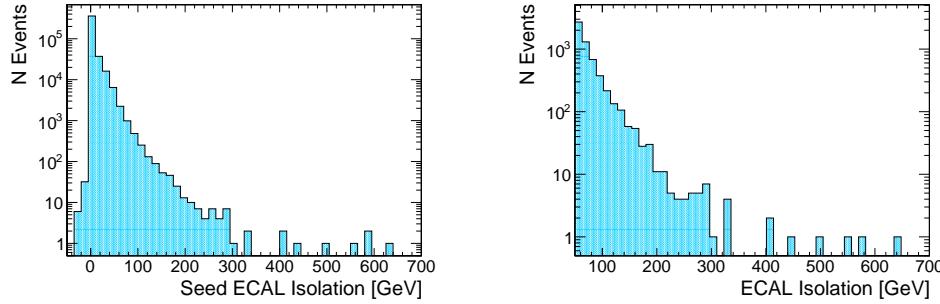
$$m_{\gamma\gamma}^2 = \Sigma^2 - \Delta^2$$

and we can conclude that so long as the splitting between the photon  $p_T$  is not too large, the scalar sum of the  $p_T$  is a good approximation to the invariant mass of a heavy object ( $m \approx \sqrt{\hat{s}}$  decay to two photons).

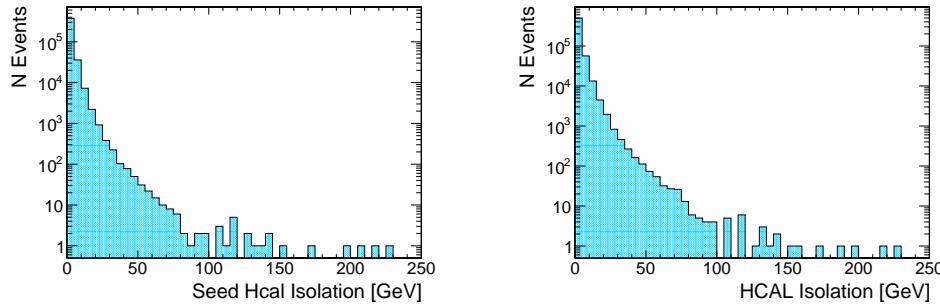
This rule of thumb has a fractional error of order  $\Delta/\Sigma$ .

### 3.3. Photon Isolation Variables.

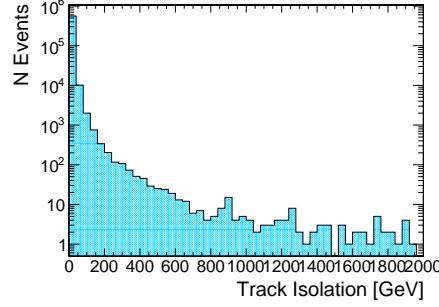
**Electromagnetic Calorimeter (ECAL) Isolation** True photons should deposit nearly all of their energy in a small cluster of crystals. To determine signal from background we sum the energy of the individual crystals hits which are not contained within some predefined area, in this case a cone size  $\Delta R > 0.4$  around the center of the cluster after subtracting out the raw energy of the reconstructed cluster. True photons will leave little to no energy outside of the cluster, whereas jet-fakes will deposit significant energy outside the cluster .



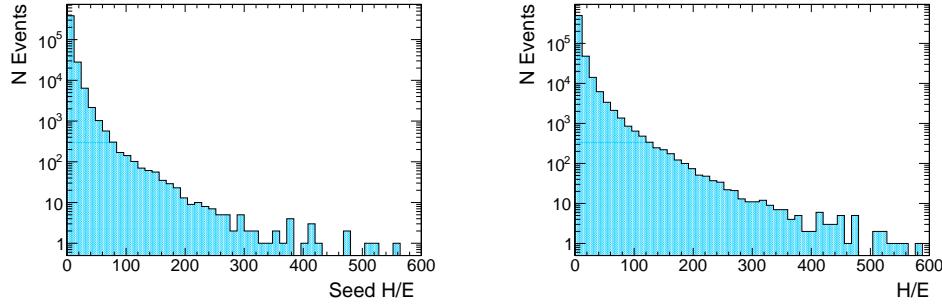
**Hadronic Calorimeter (HCAL) Isolation** Analogous to ECAL isolation, we sum all the energy in a  $\Delta R$  cone of .04 in the HCAL region behind the ECAL, but now with no subtraction for the photon. The background processes will leave relatively large deposits of energy in the HCAL.



**Tracking Isolation** The sum of the  $p_t$  of tracks within a  $\Delta R$  cone of size .03 from the center of the super-cluster. Photons do not leave tracks, but number of low  $p_t$  tracks within the cone is acceptable as nearby objects can produce tracks not associated with the photon. This quantity rejects jet fakes as jets will produce numerous tracks near the cluster. Tracking Isolation is only applied to the second leg in the high level trigger.



**H/E** The ratio of the hadronic energy of all of the HCAL hits in a cone of  $\Delta R < 0.15$  to the electromagnetic energy of the photon. This is related to the amount of hadronic showering that makes it past the ECAL rather than the distribution of energy in the HCAL



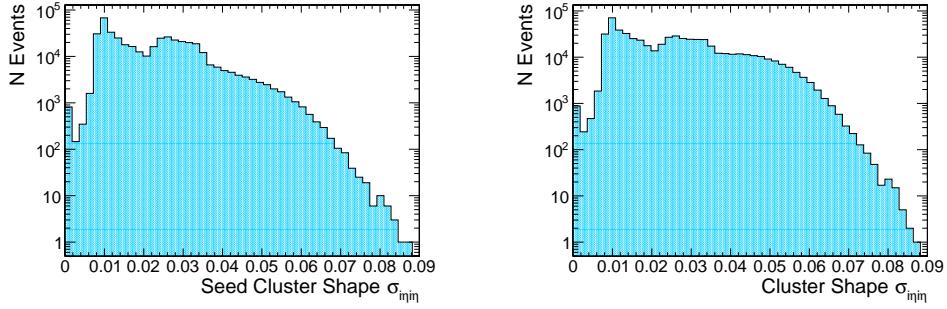
### 3.4. Photon Showershape Variables.

$\sigma_{i\eta i\eta}$  The second moment of the photon's energy cluster about its average position in pseudo-rapidity  $\eta$  weighted by a quantity  $w_i$  related to the energy deposition. Rather than using a continuous value of  $\eta$  we use a specific set of  $\eta_i$  where  $i$  is indexed by the crystals. The barrel ( $|\eta| < 1.5$ ) crystals are a different size than those installed in the endcaps so the threshold value is larger. We expect the energy to spread minimally and therefore reject clusters with large energy weighted spreading. The index  $i$  below indexes the crystals in the candidate photon cluster.

$$\sigma_{i\eta i\eta}^2 = \frac{\sum_i w_i (\eta_i - \bar{\eta})^2}{\sum_i w_i}$$

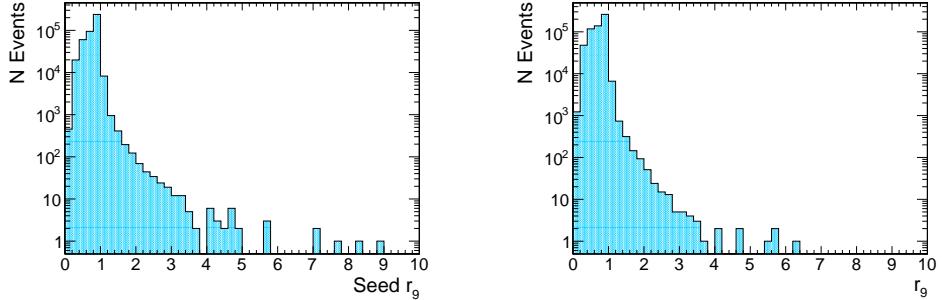
$$w_i = \max(0, 4.7 + \log(E_i/E_{5\times 5}))$$

where  $E_{5\times 5}$  is the energy of the 5 by 5 group of crystals centered on the seed crystal.



$r_9$  The ratio between the energy of the 3 by 3 array of crystals around the seed crystal and the supercluster energy. When a supercluster is generated by a single isolated photon almost all of the energy is deposited in this 3 by 3 array. In contrast, the supercluster produced by jet-fakes will have a significant portion of their energy deposited outside the 3 by 3 region as jets will tend to produce a larger spread of hits in the ECAL.

$$r_9 = \frac{E_{3 \times 3}}{E_{cluster}}$$



### 3.5. Summary Photon Identification.

Below we summarize the photon identification cuts required on both the seeded and unseeded legs of the trigger. An offset proportional to the transverse energy of the photon is applied to each of the isolations to account for an “average” amount of pileup

Category	Variable	Barrel ( $ \eta  < 1.5$ ) Cut	Endcap ( $ \eta  < 3$ ) Cut
CALO	$H/E$	< .1	< .1
CALO	$\sigma_{inj\eta}$	< .014	< .035
ISO	ECAL Isolation	$< 5 - .012 \times E_t$	$< 5 - .012 \times E_t$
ISO	HCAL Isolation	$< 5 - .005 \times E_t$	$< 5 - .005 \times E_t$
ISO	Tracking Isolation	$< 5 - .002 \times E_t$	$< 5 - .002 \times E_t$
R9	$r_9$	> .85	> .85

where both photons in the two photon trigger are required to either pass the logical bit: (R9 OR (ISO AND CALO)). These cuts are designed to mimic the offline cuts and categories use for the  $H \rightarrow \gamma\gamma$  analysis.

## 4. HLT DEFINITIONS

### The Level 1 (L1) Trigger Bits

The high level trigger must be seeded by the L1 trigger. The most important to this study being:

- (1) L1\_DoubleEG\_13\_7: Requires two ECAL objects of transverse momentum greater than 13 and 7 GeV.
- (2) L1\_SingleEG\_22: Requires a single ECAL object of transverse momentum greater than 22 GeV.

Here “EG” stands for egamma (electron photon).

### The High Level Trigger (HLT) Bits

The  $\mathcal{L} = 8 \times 10^{33} \text{ cm}^{-2}s^{-1}$  trigger suite consisting of 3 bits, each seeded by a level 1 trigger. Each bit requires two photons with the same photon identification criterion, but differing kinematic requirements.

Knowing now that  $m_H \approx 125 \text{ GeV}$ , the most straightforward way to reduce the rate of unwanted events is by placing a cut on the invariant mass of the two photons while keeping the transverse energy  $E_T$  as low as possible. However, as the  $H \rightarrow \gamma\gamma$  is not the only analysis which uses these triggers, there is a need for triggers with no invariant mass requirement (typically in non-resonant diphoton searches). As the mass70 Trigger accounts for most of the  $H \rightarrow \gamma\gamma$  budget, the  $E_T$  requirements on the no mass cut trigger are significantly higher.

- (1) 26/18 with a 70 GeV mass cut seeded by L1\_DoubleEG\_13\_7:  
`HLT_Photon26_R9Id85_OR_CaloId10_Iso50_Photon18_R9Id85_OR_CaloId10_Iso50_Mass70_v2`
- (2) 36/18 with no mass cut seeded by L1\_SingleEG\_22:  
`HLT_Photon36_R9Id85_OR_CaloId10_Iso50_Photon22_R9Id85_OR_CaloId10_Iso50_v6`
- (3) 36/10 with a mass 80 cut seeded by L1\_SingleEG\_22:  
`HLT_Photon36_R9Id85_OR_CaloId10_Iso50_Photon10_R9Id85_OR_CaloId10_Iso50_Mass80_v1`

As the level 1 triggers are not fully efficient at their nominal value, the HLT cuts are placed roughly  $2 \times$  the thresholds of the L1 triggers. For instance the HLT 26/18 is roughly twice the DoubleL1 13/7 and the HLT 36/18 is roughly twice the Single L1 22. The triggers are kept as asymmetric as possible to remain efficient while rejecting processes which favor balanced  $p_T$ .

Historically, all of the triggers were seeded by `SingleEG` L1 triggers (although with different thresholds), but the rate could no longer be maintained and the highest rate trigger, the mass70 was moved to a double photon level 1 trigger. The mass 70 trigger now identifies the first leg of the trigger as the seeded leg and then second leg as the unseeded leg, although the name has lost significance when the L1 dependance was altered.

The first two triggers are the main source of events whereas the third trigger has a large overlap and can generally be ignored insignificant to the cumulative rate calculation.

## 5. ANALYSIS CONSIDERATIONS

Now that we have a clear understanding of what the triggers are designed to do and the parameter space in which they are defined we can address issues related to the analysis workflow.

The first issues I will address are practical issues related to the rate study, but later I will motivate ways the  $H \rightarrow \gamma\gamma$  analysis can be changed to reduce the rate for higher luminosity running.

### 5.1. OpenHLT Simulation and Rate Analysis.

Historically, the CMS Trigger group developed a tool known as OpenHLT, designed to allow fast simulation of triggers on RAW data. As part of the OpenHLT framework, analysis code was available to quickly perform rate calculations for any given trigger suite. Unfortunately, since the latest version of the CMS software package (**CMSSW**) the CMS Trigger group has “encouraged” groups to develop their own trigger simulation analyzers. The change was motivated by the fact that, analysis groups frequently needed to create their own analyzers on top of the OpenHLT framework.

The CMS Trigger group did provide an update up the OpenHLT tool to work with new versions of **CMSSW**, which I will refer to as “new” OpenHLT. The new tool has no accompanying analyzers to perform rate calculations. The tool is capable of simulating a modified trigger menu, but outputs files excessively large  $O(10 \text{ Gbs})$  per simulated trigger for performing trigger parameter scans in a overly complicated file format of which Trigger Experts are the main users. The tool was designed for someone trying to simulate a single trigger quickly and check its efficiency. However, for our purposes we would like to scan a large area of parameter space and access the HLT level photon identification variables and kinematics easily.

To deal with the limitations of this tool, I developed scripts to run the “new” OpenHLT tool, parse relevant information for the trigger scan, promptly delete the unnecessary information, and store the information in a lightweight, easily accessible tree format.

Code was also developed to perform the event by event trigger logic for every trigger menu in space of the 3 diphoton triggers scanned in the leading and subleading photon  $p_T$ .

**5.2. Tag and Probe.** Currently the Higgs analysis uses a mass cut of 70 on it’s “workhorse” trigger to keep the analysis within the  $Z$  boson peak  $m_Z \approx 91 \text{ GeV}$ . This is important for the analysis as a candle for finding good electrons. The procedure for this is referred to as “tag and probe”.

Photons and electrons are similar enough in character that we can consider them as “good” electromagnetic energy and select good EM energy in the same way, save a couple exceptions. Electrons, as they are charged, will leave a track in the tracker and register a pixel seed in the pixel detector.

The tag and probe procedure for gathering an unbiased sample of quality EM objects from the  $Z$  peak is as follows:

- An event must first pass the 26/18 HLT trigger with minimal requirements on isolation for the first leg and only basic requirements on the second leg i.e. the existence of a moderately energetic ECAL cluster.
- The event must pass the offline analysis cuts made excluding the pixel seed veto (designed to reject electrons)
- The invariant mass of the two photons must be compatible with the  $Z$  peak (between 70 and 110 GeV). This is what sets the bound on the trigger.
- The first leg is “tagged” by the trigger decision and is thus biased, so we use the second leg of the trig as the “probe” for a given study. The probe is a nearly unbiased sample of good electromagnetic objects coming from a  $Z \rightarrow ee$  decay.

This procedure is an important tool in the analysis, but  $H \rightarrow \gamma\gamma$  is not the only analysis performing tag and probe with  $Z \rightarrow ee$  decays. These events are kept by single and double lepton HLT triggers. The analysis could foreseeably move the mass cut of the double photon trigger above the  $Z$  peak to 100 GeV for

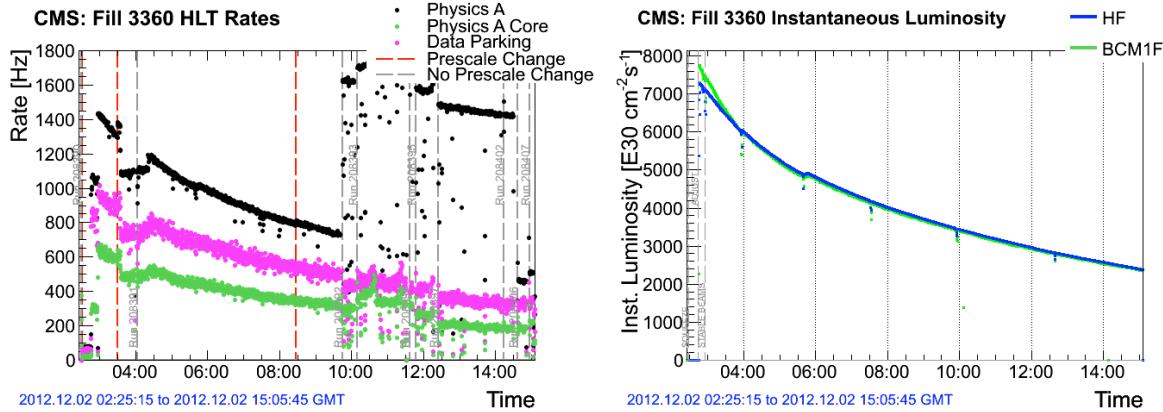


FIGURE 7. (left) the HLT rate as a function of time for fill 3360. As you can see run 208390 was relatively stable to the rest of the run. (right) the instantaneous luminosity of the beams as a function of time. The study was performed at the highest stable luminosity possible in preparation for higher luminosity in 2015.

a significant reduction in rate. The background is not even fit below 100 GeV as shown in Figure 1. The analysis would need to process the lepton samples, but this inconvenience is secondary to the bandwidth restrictions as the luminosity continues to increase.

### 5.3. Sample Run 208390.

In preparation for the higher luminosity which will accompany the higher energy collisions in 2015 we perform trigger scans in the a recent stable run of  $\sqrt{s} = 8$  TeV, Run 208390 (Fill 3360), containing  $25.86 \text{ pb}^{-1}$  of data. This specific run was chosen for its rate stability, relatively large integrated luminosity, and high instantaneous luminosity within the  $\sqrt{s} = 8$  TeV program (Figure 7).

Within a given suite of triggers we define the following quantities.

**Unique Rate** The rate of the trigger corresponding to events which only fire said trigger. Expressed logically for trigger bits  $t_1, \dots, t_N$  we have:  $t_1 \text{ AND } !t_2 \text{ AND } \dots \text{ ! } t_N$

**Pure Rate** The rate of the trigger caused by the addition of a trigger to a suite of already existing triggers. As the order in which you consider in the triggers can change this value, it is not unique. For instance say trigger #1 triggers on 90/100 events and trigger #2 triggers on 95/100 events, but all 90 of the first trigger's events were included. The pure rate of adding trigger #1 would be  $90k$  and the the pure rate of adding trigger #2 would be  $5k$  where  $k$  is some scaling constant between events and Hz. However, if we had considered adding trigger #2 first it would have a pure rate of  $95k$  and the adding trigger #1 would have 0 pure rate. This is used as a measure of the costliness of adding a new trigger to the menu.

The pure and the unique rate are the most difficult rates to simulate as they require an **EVENT by EVENT** calculation of the bit logic. Suppose you wish to analyze a trigger suite of  $N_t$  triggers, for each trigger you would like to scan  $p$  points where at which point you have different parameters of the trigger. Then the number of calculations would scale as a power law:

$$N_{\text{calc}} = N_{\text{events}} \times p^{N_t} = 10^4 \times (10^2)^2 \approx 10^8$$

To estimate how long it would take to perform this scan we must scale by the time it takes to simulate a single event of the HLT and compute the bit logic.

**Raw Rate** The rate of the trigger corresponding to all events in which the trigger fired, regardless of the other triggers in the suite. This number is universal and is quoted for every trigger in the report generated

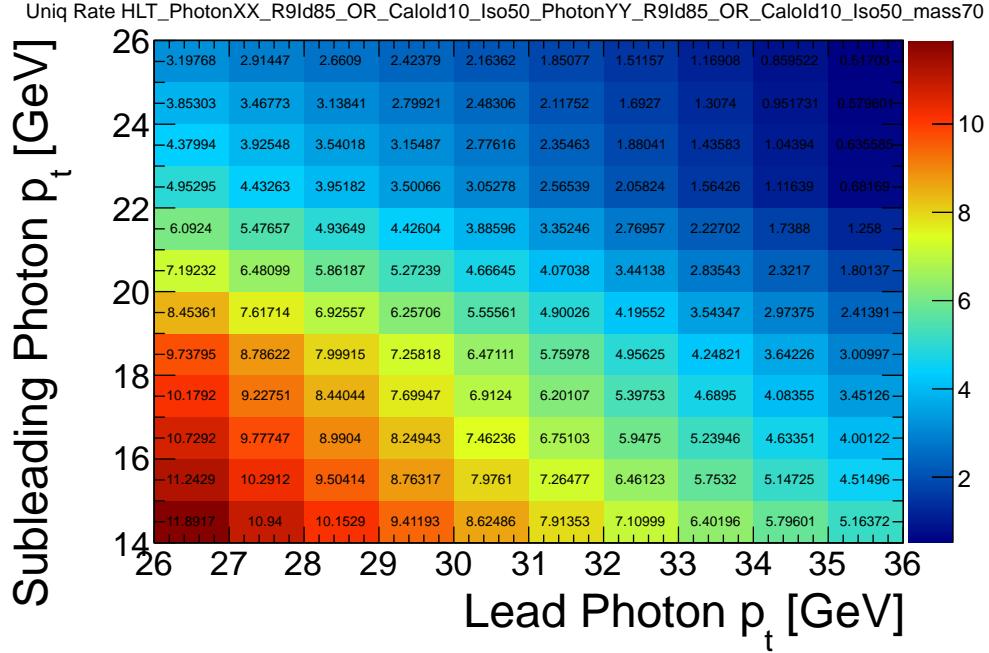


FIGURE 8. The unique rate of the mass 70 trigger fixing the no mass cut trigger

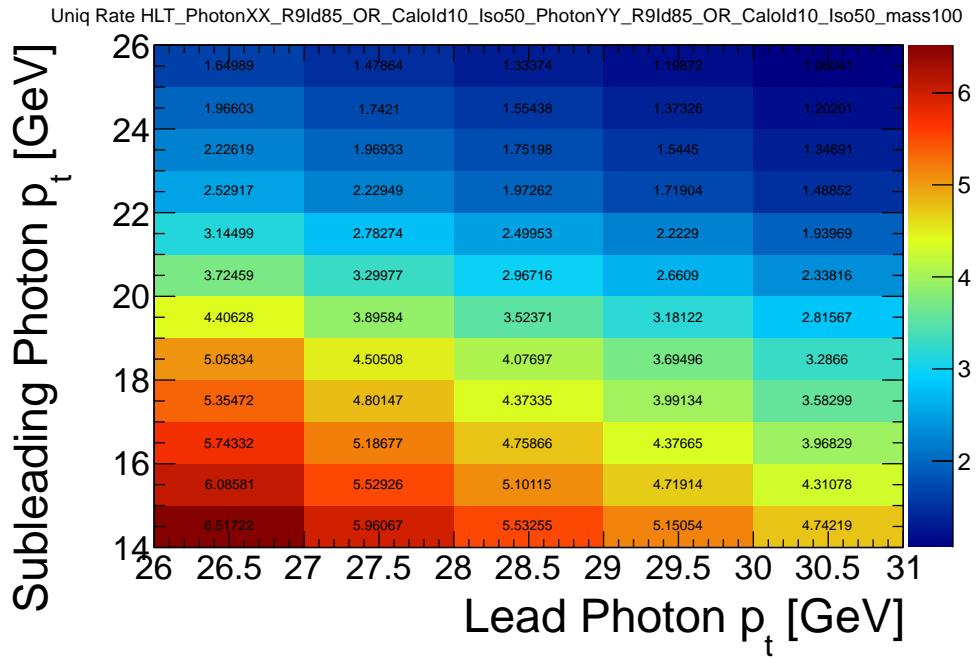


FIGURE 9. The unique rate of the mass 100 trigger fixing the no mass cut trigger

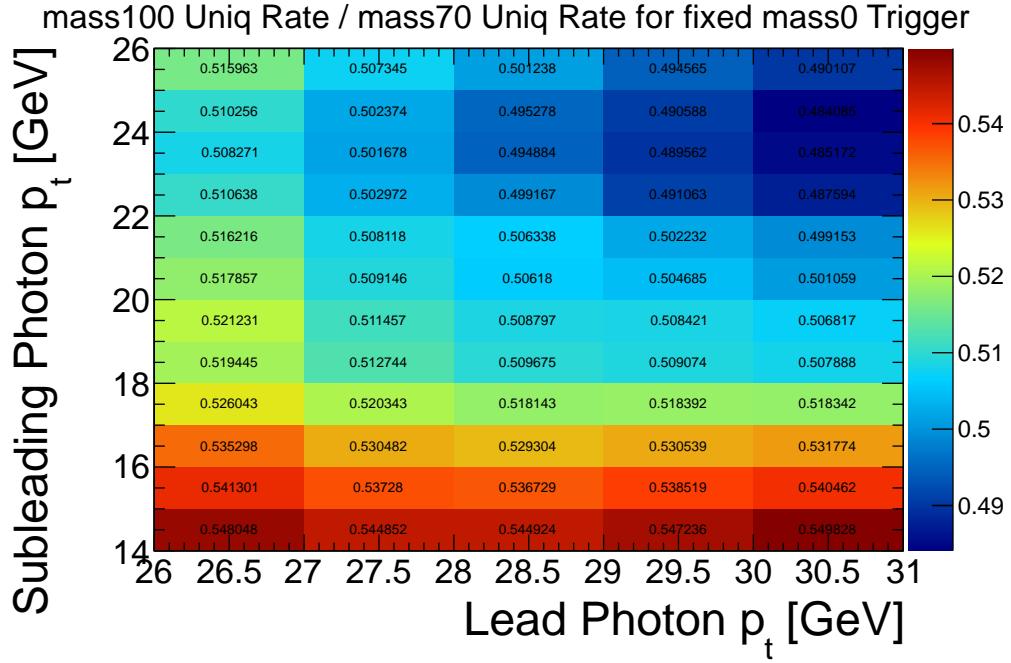


FIGURE 10. The ratio of the mass 70 and mass 100 unique rates fixing the no mass cut trigger

after every run. This rate is not incredibly useful for measuring suites of triggers that have a large overlap in events they will trigger on.

**Cumulative Rate** The actual rate for the trigger suite. This is the rate that must be kept within the budget. It is the rate corresponding to the events which pass any of the triggers in the suite. Expressed logically:  $t_1 \text{ OR } t_2 \text{ OR } t_3 \dots \text{ OR } t_N$

## 6. CONCLUSIONS

## REFERENCES

PRINCETON UNIVERSITY, DEPARTMENT OF PHYSICS JADWIN HALL, PRINCETON NJ 08544

*E-mail address:* joshuarh@princeton.edu

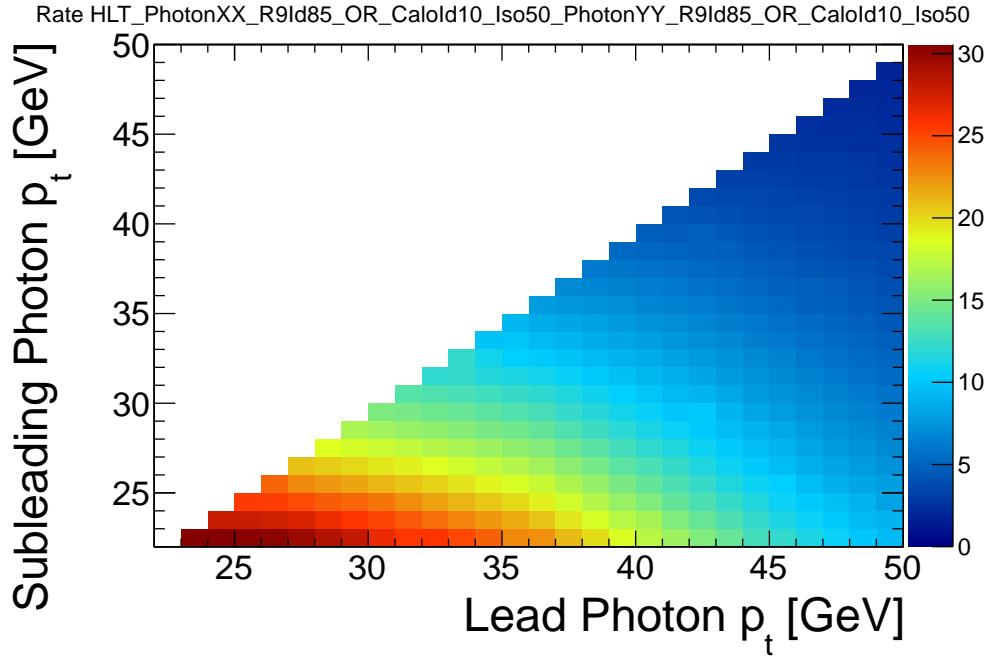


FIGURE 11. The raw rate of the no mass cut trigger as a function of thresholds on the lead and subleading photon  $p_t$  threshold

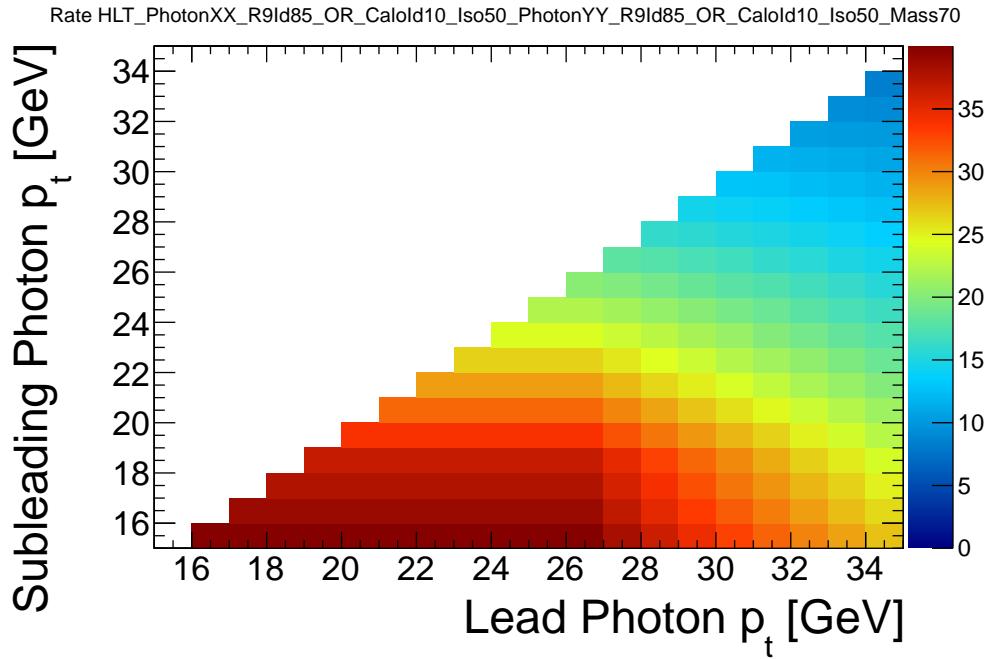


FIGURE 12. The raw rate of the mass 70 cut trigger as a function of the thresholds on the lead and subleading photon  $p_t$  threshold