

# Search for the Flavor Changing Neutral Current, $t \rightarrow q\gamma$ in Top Pair Events Using the ATLAS Detector

Jason Barkeloo

April 16, 2020



# Welcome!

Thank you for your flexibility given our current situation!

Defense proceedings:

- ▶ Open/public presentation in this room with followup questions
- ▶ Committee and myself will adjourn to a separate room to conclude
- ▶ Committee will deliberate and I will return to this room

General Guidelines:

- ▶ Please mute if not speaking

**Thank you for attending!**

# Overview

Background

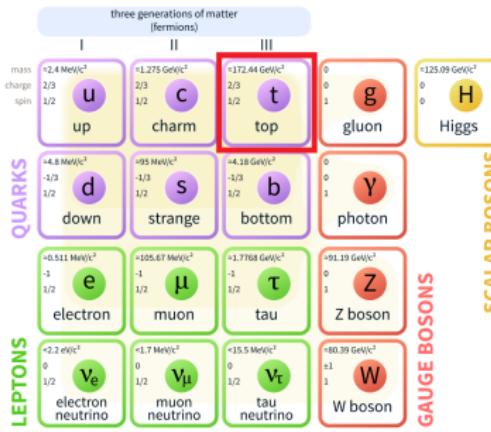
The LHC and ATLAS

Searching for Flavor Changing Neutral Current Signatures

Conclusion

# The Standard Model

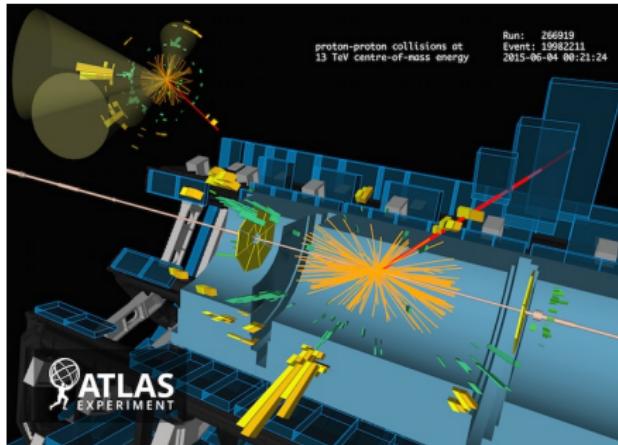
## Standard Model of Elementary Particles



- Best working description of fundamental particles and their interactions
  - Experimentally precise and well behaved

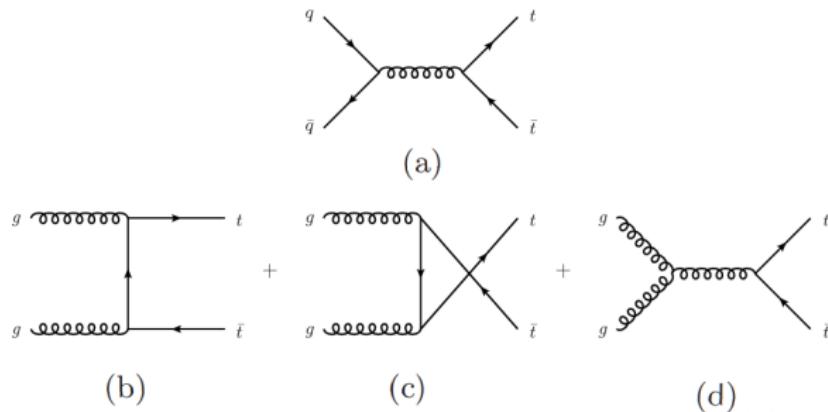
# The Top Quark

- ▶ Heaviest fundamental particle,  $172.5\text{ GeV}$
- ▶ Lifetime  $5 \times 10^{-25}\text{ s}$ , decays before hadronization
  - ▶ Allows us to study the decay of a single quark



# Top Quark Pair Production

- ▶ Leading order processes for top quark production
  - ▶ Quark-antiquark annihilation  $\approx 10\%$
  - ▶ Gluon-gluon fusion  $\approx 90\%$



**Figure:** Leading order  $t\bar{t}$  diagrams

# Top Quark Pair Production

- At  $\sqrt{s} = 13 \text{ TeV}$  for  $m_t = 172.5 \text{ GeV}$ ,  $\sigma_{t\bar{t}} = 831.76 \text{ pb}$

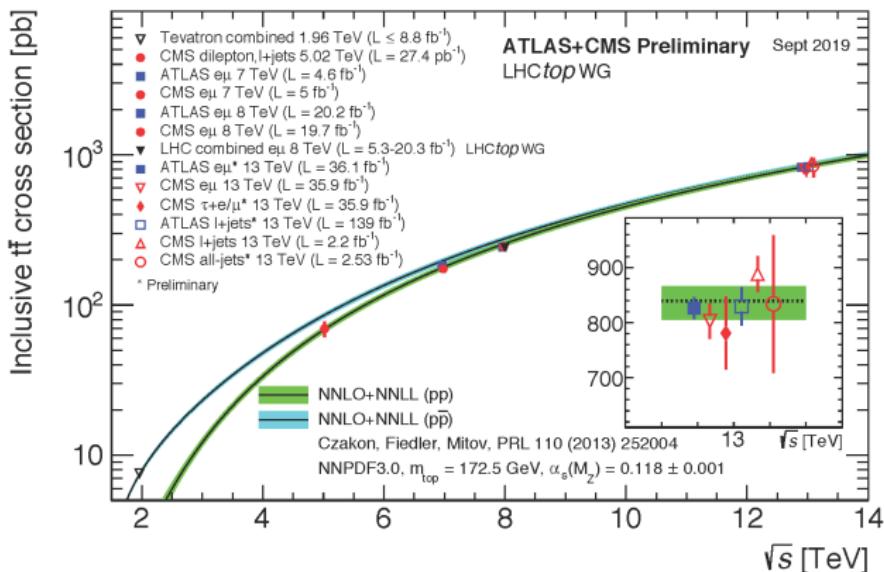
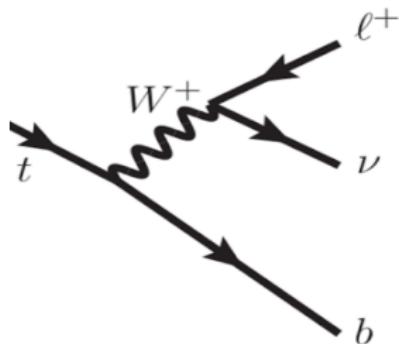


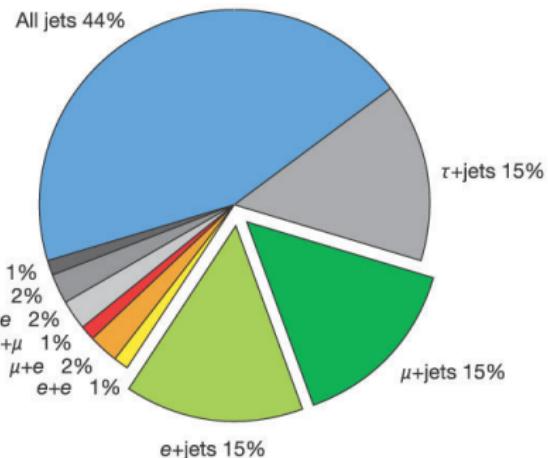
Figure:  $t\bar{t}$  production cross section [TopWGSummaryPlots]

# Top Quark Decays

- ▶ Standard model top branching ratio to  $bW \simeq 100\%$

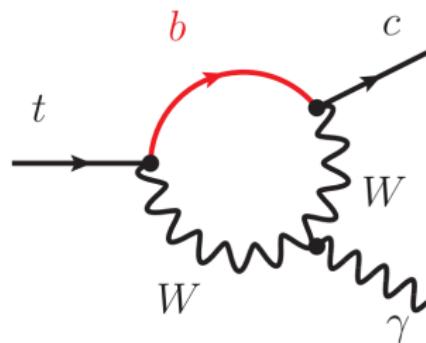
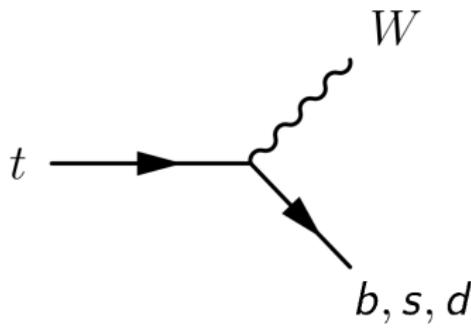


**Figure:** Leptonic final state diagram for a top decay



**Figure:** Top quark pair decay final states [Nature]

# Top Quark Decays in the SM



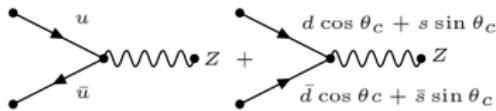
- ▶  $t \rightarrow bW \approx 99.83\%$
- ▶  $t \rightarrow sW \approx 0.16\%$
- ▶  $t \rightarrow dW \approx 0.01\%$

- ▶  $t \rightarrow q_{u,c} X \approx 10^{-17} - 10^{-12}$
- ▶ Limits on  $t \rightarrow \gamma q$  processes:  
[Phys.Lett. B800 135082]
  - ▶  $t \rightarrow \gamma u < 2.8 \times 10^{-5}$
  - ▶  $t \rightarrow \gamma c < 18 \times 10^{-5}$

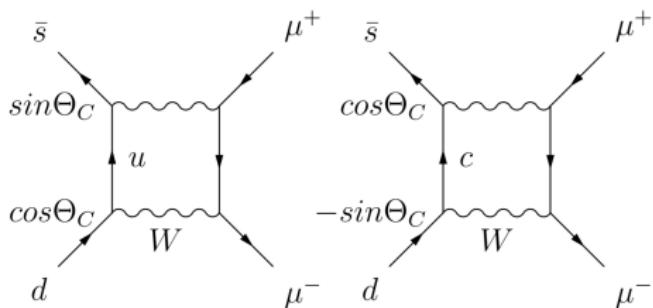
# GIM Mechanism

- ▶ Cabibbo model - 3 quarks (u, d, s)
- ▶ Studies of kaon decays showed the existence of  $K^+ \rightarrow \mu^+ \nu_\mu$  but an absence of predicted  $K_L^0 \rightarrow \mu^+ \mu^-$
- ▶ Even in the absence of a tree level decay  $K_L^0$  decay the box diagram would be possible through an exchange of W bosons
- ▶ Weak neutral current interactions in the uds model have the form

$$u\bar{u} + (d\bar{d} \cos^2 \theta_C + s\bar{s} \sin^2 \theta_C) + (s\bar{d} + d\bar{s}) \sin \theta_C \cos \theta_C$$



# GIM Mechanism

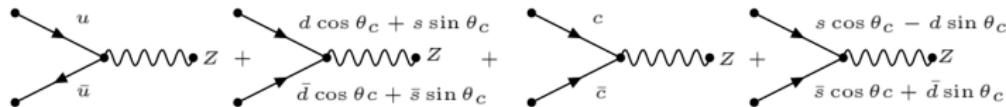


- ▶ Glashow, Iliopoulos, and Maiani [Phys. Rev. D (1970)] propose a mechanism through which FCNCs are suppressed in loop diagrams
  - ▶ Introduction of charm quark
- ▶ Kaon decays imply no neutral current/natural suppression of neutral current

# GIM Mechanism

$$\begin{bmatrix} d' \\ s' \end{bmatrix} = \begin{bmatrix} \cos \theta_c & \sin \theta_c \\ -\sin \theta_c & \cos \theta_c \end{bmatrix} \begin{bmatrix} d \\ s \end{bmatrix}$$

- ▶ The addition of the charm changes our weak neutral current interactions
  - ▶ With four quarks the weak neutral interactions now have the form:
- $$u\bar{u} + c\bar{c} + (d\bar{d} + s\bar{s}) \cos^2 \theta_C + (s\bar{s} + d\bar{d}) \sin^2 \theta_C + (s\bar{d} + d\bar{s} - d\bar{s} - s\bar{d}) \sin \theta_C \cos \theta_C$$
- ▶ Flavor changing neutral current diagrams cancel out at tree level (as  $m_c \rightarrow m_u$ )



# CKM Matrix

$$\begin{bmatrix} |V_{ud}| & |V_{us}| & |V_{ub}| \\ |V_{cd}| & |V_{cs}| & |V_{cb}| \\ |V_{td}| & |V_{ts}| & |V_{tb}| \end{bmatrix} = \begin{bmatrix} 0.97427 \pm 0.00015 & 0.22534 \pm 0.00065 & 0.00351^{+0.00015}_{-0.00014} \\ 0.22520 \pm 0.00065 & 0.97344 \pm 0.00016 & 0.0412^{+0.0011}_{-0.0005} \\ 0.00867^{+0.00029}_{-0.00031} & 0.0404^{+0.0011}_{-0.0005} & 0.999146^{+0.000021}_{-0.000046} \end{bmatrix}$$

Figure: CKM Matrix

- ▶ Decay rates proportional to  $|V_{tx}|^2$
- ▶ Top decay through a  $W^\pm$  boson is a charged current interaction.
- ▶ Flavor changing processes are proportional to off-diagonal elements of the CKM matrix
- ▶ GIM/CKM suppression of these FCNC processes in the Standard Model make them unlikely to be seen without some new physics

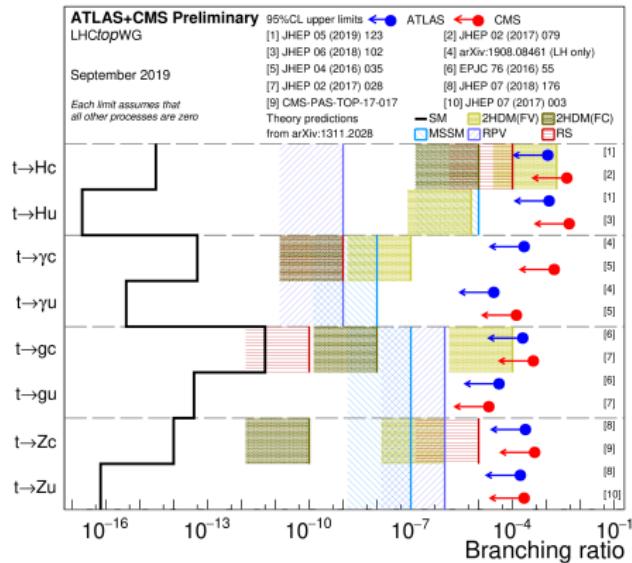
# Top Flavor Changing Neutral Currents (FCNCs)

Process	SM	2HDM(FV)	2HDM(FC)	MSSM	RPV	RS
$t \rightarrow Zu$	$7 \times 10^{-17}$	—	—	$\leq 10^{-7}$	$\leq 10^{-6}$	—
$t \rightarrow Zc$	$1 \times 10^{-14}$	$\leq 10^{-6}$	$\leq 10^{-10}$	$\leq 10^{-7}$	$\leq 10^{-6}$	$\leq 10^{-5}$
$t \rightarrow gu$	$4 \times 10^{-14}$	—	—	$\leq 10^{-7}$	$\leq 10^{-6}$	—
$t \rightarrow gc$	$5 \times 10^{-12}$	$\leq 10^{-4}$	$\leq 10^{-8}$	$\leq 10^{-7}$	$\leq 10^{-6}$	$\leq 10^{-10}$
$t \rightarrow \gamma u$	$4 \times 10^{-16}$	—	—	$\leq 10^{-8}$	$\leq 10^{-9}$	—
$t \rightarrow \gamma c$	$5 \times 10^{-14}$	$\leq 10^{-7}$	$\leq 10^{-9}$	$\leq 10^{-8}$	$\leq 10^{-9}$	$\leq 10^{-9}$
$t \rightarrow hu$	$2 \times 10^{-17}$	$6 \times 10^{-6}$	—	$\leq 10^{-5}$	$\leq 10^{-9}$	—
$t \rightarrow hc$	$3 \times 10^{-15}$	$2 \times 10^{-3}$	$\leq 10^{-5}$	$\leq 10^{-5}$	$\leq 10^{-9}$	$\leq 10^{-4}$

**Table:** Branching ratio enhancements in various beyond the standard model theories [Snowmass Top Report]

# Top Flavor Changing Neutral Currents

## ► Current Limits on FCNC Decays



## ► Limits on $t \rightarrow q\gamma$ processes: [JHEP 04 (2016) 035]

- $t \rightarrow \gamma u < 2.8 \times 10^{-5}$
- $t \rightarrow \gamma c < 18 \times 10^{-5}$

# Monte Carlo Production of FCNC Signal Samples

- ▶ Due to the low cross sections we must create our own Monte Carlo Samples for our Signal
- ▶ An effective field theory approach was taken in the creation of the model [Degrande et al. Phys. Rev. D 91, 034024 (2015)]
- ▶ This model takes advantage of dimension-6 operators

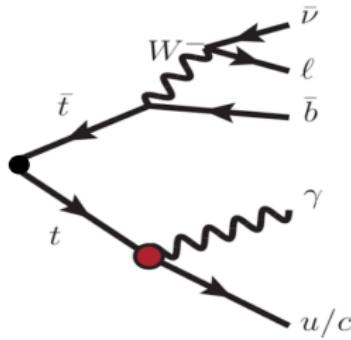
$$\mathcal{L}_{SM} = \mathcal{L}_{SM}^{(4)} + \mathcal{L}^{eff} \text{ where } \mathcal{L}^{eff} = \frac{1}{\Lambda^2} \sum_k C_k^{(6)} Q_k^{(6)}$$

$$\mathcal{L}_{tq\gamma}^{eff} = C \sigma^{\mu\nu} q_\nu (\lambda_{ct}^L P_L + \lambda_{ct}^R P_R) t A_\mu + H.c.$$

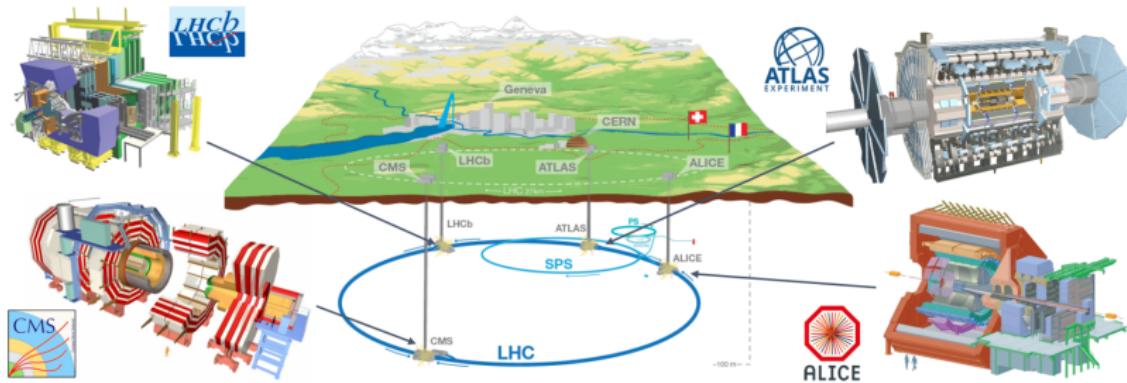
# FCNC: What are we looking for? $t\bar{t} \rightarrow W(\rightarrow l\nu) b + q\gamma$

- ▶ Final state topology

- ▶ One Neutrino, from W
- ▶ One Lepton, from W
- ▶ One B-jet, SM Top
- ▶ One Photon, FCNC Top
- ▶ One Jet, FCNC Top

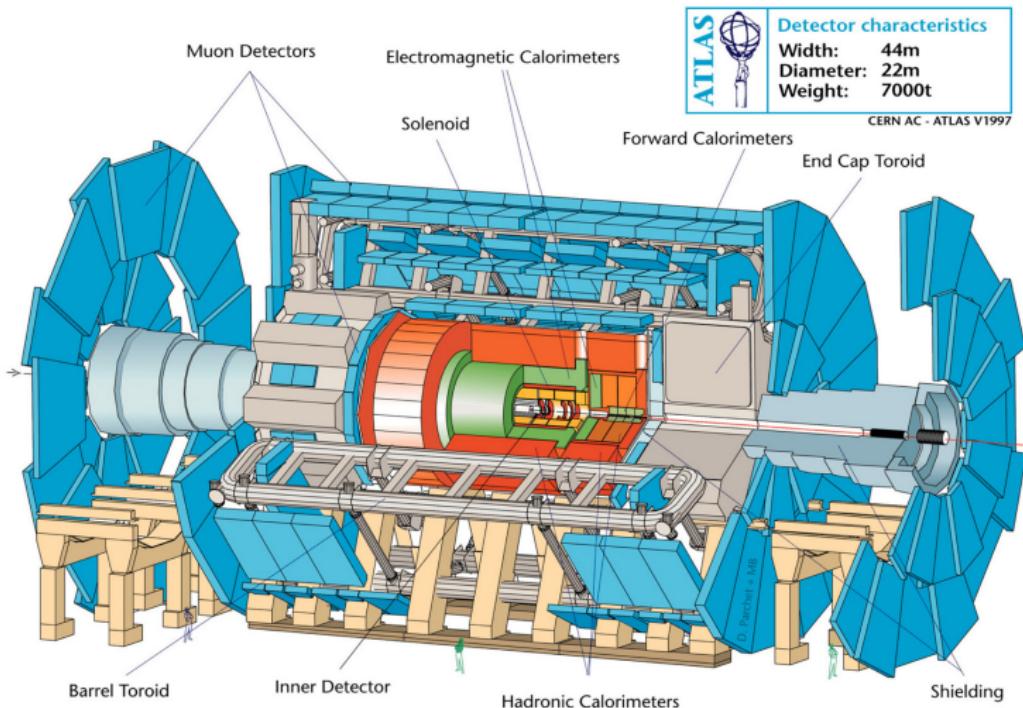


# The Large Hadron Collider

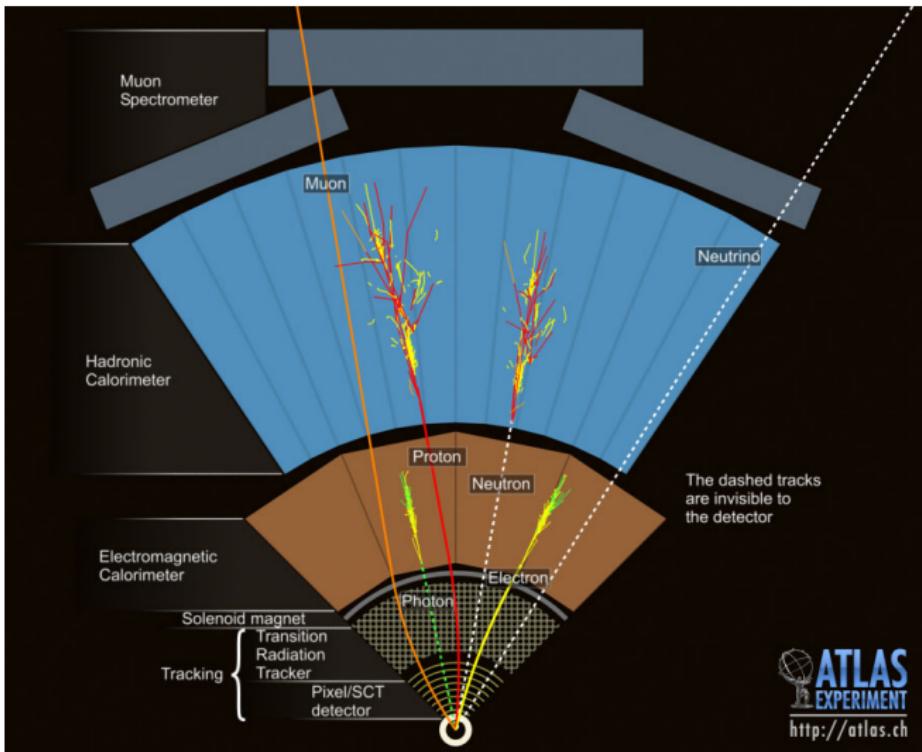


- ▶ 27km ring beneath Franco-Swiss border
- ▶ Collides protons at center of mass energy 13TeV
- ▶ Over 10 Quadrillion ( $10^{15}$ ) events produced within the ATLAS detector so far

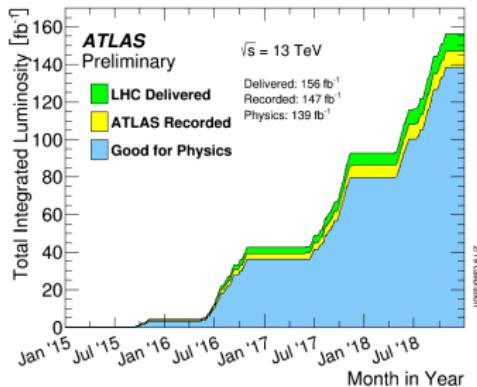
# The ATLAS Detector



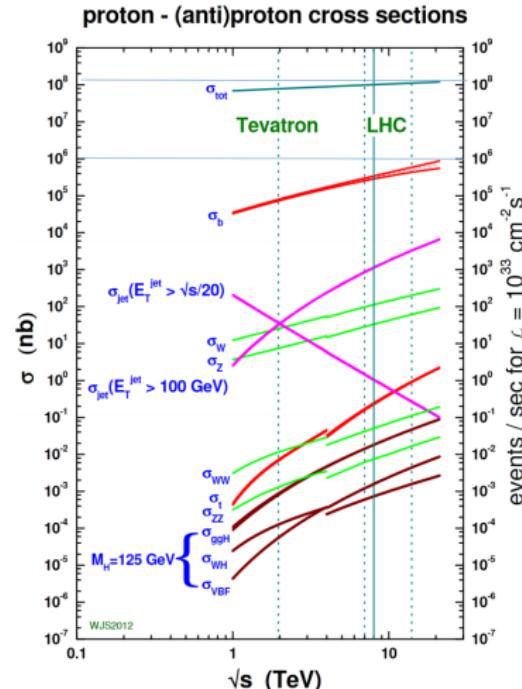
# Particles in ATLAS



# ATLAS Data

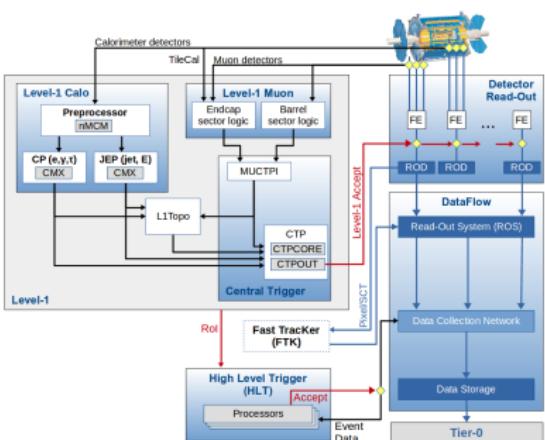
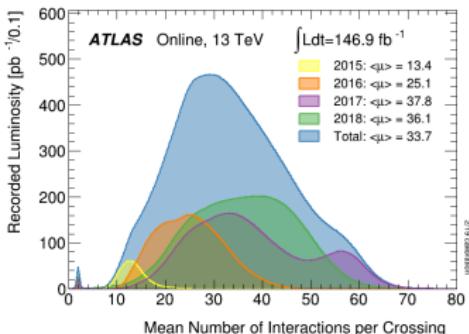


- ▶ Total Luminosity at  $\sqrt{s} = 13 \text{ TeV}$ :  $139 \text{ fb}^{-1}$
- ▶ The number of events we see is  $N = \sigma L$
- ▶  $N_{tot} \approx 14 \times 10^{15}$  events produced during the 13TeV data runs
- ▶  $N_{t\bar{t}} \approx 116 \times 10^6$



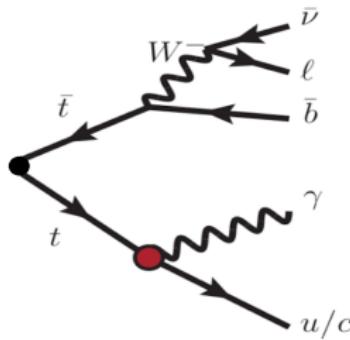
# ATLAS Data Acquisition

- ▶ Average 33.7 collisions per crossing during Run 2, 40MHz collision rate
- ▶ Raw uncompressed event: 1.6 MB
- ▶ Around 64 TB/s data
- ▶ 2 Stage Trigger System
  - ▶ Level 1 - Hardware based, coarse object reconstruction, reduces rate to under 100 kHz
  - ▶ High-Level Trigger: Software based, performs reconstruction as close to offline as possible, reduces rate to 1 kHz
- ▶ Still save 10s of PB per year



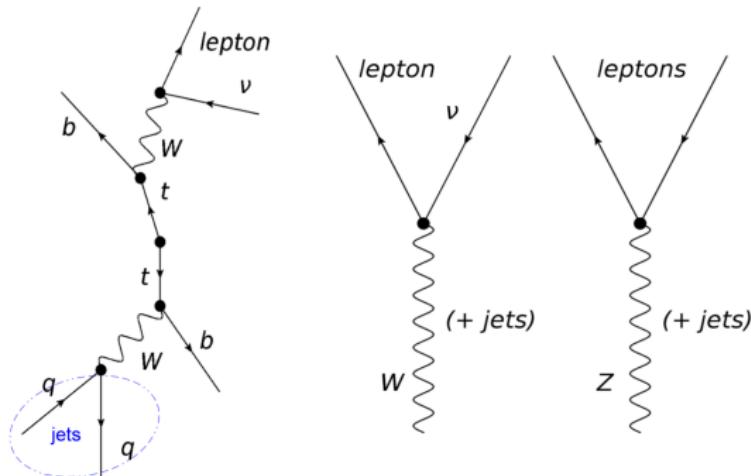
# Object Preselection

- ▶ We preselect events with objects that look like similar to our expected topology
- ▶ Require:
  - ▶ Exactly one lepton ( $e$  or  $\mu$ )  $\geq 25$  GeV
  - ▶ Exactly one good photon  $\geq 15$  GeV
  - ▶ Missing Transverse Energy  $\geq 30$  GeV
  - ▶  $\geq 1$  Jets



# Background Processes

- ▶ Due to all of the processes at hadron colliders it is important to model similar event topologies well.
- ▶ Major backgrounds include  $t\bar{t}$ , W+Jets, Z+Jets, + processes with an associated photon



# 0 Photon Backgrounds

# Real Photon Backgrounds

# Fake Rate Object Selection

- ▶ Want to calculate fake rate in events which could enter the signal region.
- ▶ Create 2 control regions:  $Z \rightarrow ee$  and  $Z \rightarrow e\gamma$
- ▶ Require:
  - ▶ Common Object Selection (MET, Jets, Triggers, etc.)
  - ▶  $Z \rightarrow ee$  : 2 Opposite Sign Electrons,  $86.1 \text{ GeV} < m_{e^+e^-} < 96.1 \text{ GeV}$
  - ▶  $Z \rightarrow e\gamma$  : 1 Electron,  $\geq 1$  Photon,  $86.1 \text{ GeV} < m_{e\gamma} < 96.1 \text{ GeV}$
- ▶ Tag and Probe Method used

# Scale Factor

$$\text{FR}^{\text{e-fake}} = \frac{N_{e,\gamma}}{N_{e,e}}$$

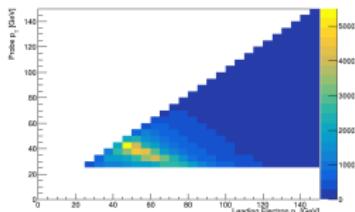
$$SF_{\text{FR}}^{\text{e-fake}} = \frac{\text{FR}_{\text{data}}^{\text{e-fake}}}{\text{FR}_{\text{MC}}^{\text{e-fake}}}$$

This scale factor is calculated for converted and unconverted photons as well as in bins of  $\eta$  and  $\phi$

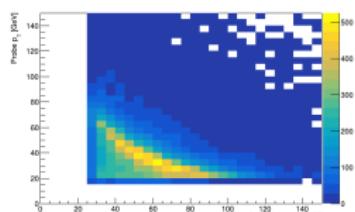
- ▶ Converted photons pair produce before the ECAL leaving tracks in the Inner Detector
- ▶ Unconverted photons only pair produce inside of the ECAL

# Data and MC Distributions

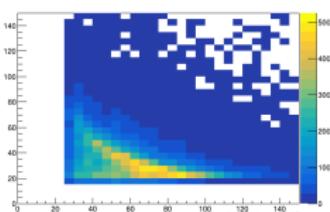
Data



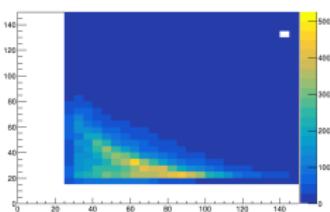
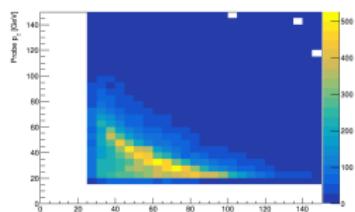
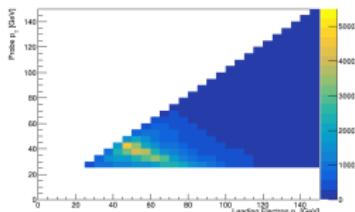
► Converted  $\gamma$



► Unconverted  $\gamma$

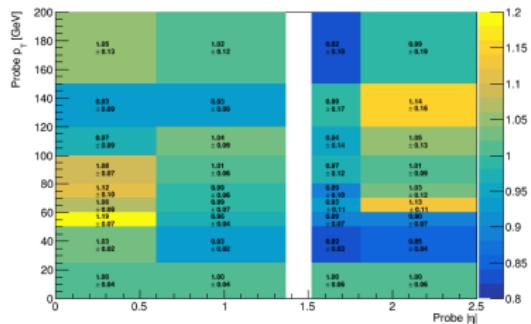


MC

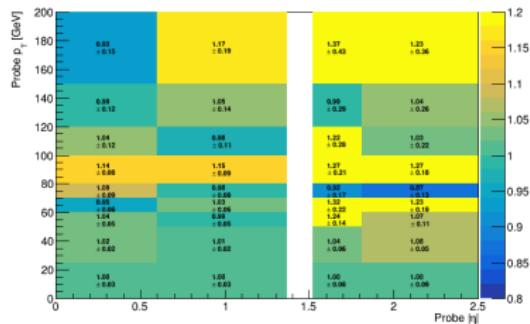


# 2D Fake Rates

► Converted  $\gamma$

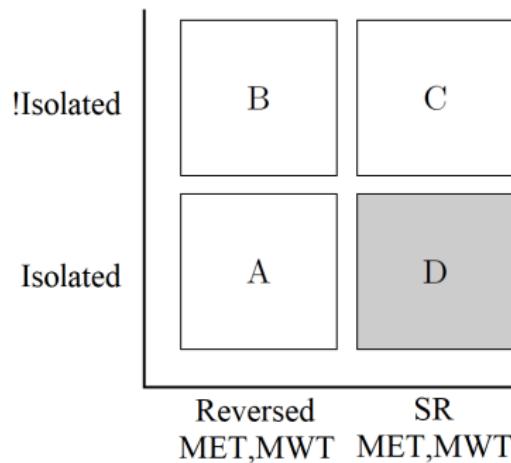


► Unconverted  $\gamma$



# $j \rightarrow \gamma$ Fake Rate Studies

Majority of hadronic fake photons are from  $t\bar{t}$  events where a final state jet radiates a non-prompt photon. Similarly radiated photons for W+jets and single top processes can enter the signal region through the radiation of a non-prompt photon.



# ABCD Method

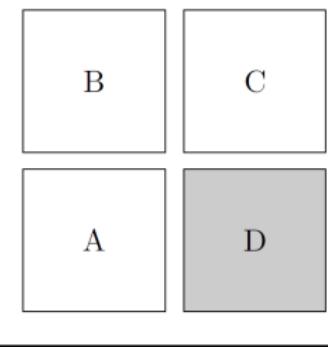
$$\frac{N_D^{\text{h-fake}}}{N_C^{\text{h-fake}}} = \frac{N_A^{\text{h-fake}}}{N_B^{\text{h-fake}}} \text{ and } \frac{N_D^{\text{h-fake}}}{N_A^{\text{h-fake}}} = \frac{N_C^{\text{h-fake}}}{N_B^{\text{h-fake}}}$$

!Isolated

Want uncorrelated variables, use a correction factor to account to ensure closure

$$\theta_{\text{MC}} = \frac{N_{D,\text{MC}}^{\text{h-fake}} / N_{C,\text{MC}}^{\text{h-fake}}}{N_{A,\text{MC}}^{\text{h-fake}} / N_{B,\text{MC}}^{\text{h-fake}}}$$

Isolated



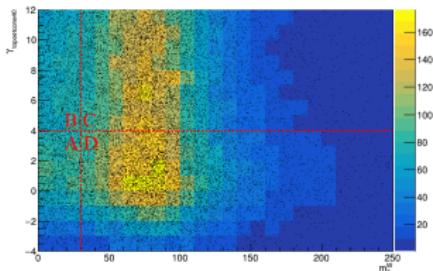
Reversed SR  
MET,MWT MET,MWT

$$N_{D,\text{est.}}^{\text{h-fake}} = \frac{N_{A,\text{data}}^{\text{h-fake}} \times N_{C,\text{data}}^{\text{h-fake}}}{N_{B,\text{data}}^{\text{h-fake}}} \times \theta_{\text{MC}}$$

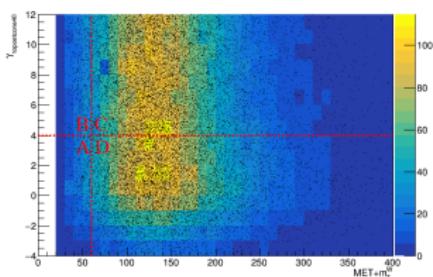
$$\text{SF}^{\text{h-fake}} = \frac{N_{D,\text{est.}}^{\text{h-fake}}}{N_{D,\text{MC}}^{\text{h-fake}}}$$

## ► Converted Photons

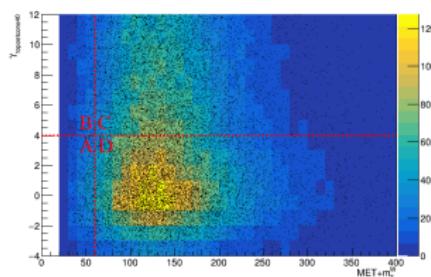
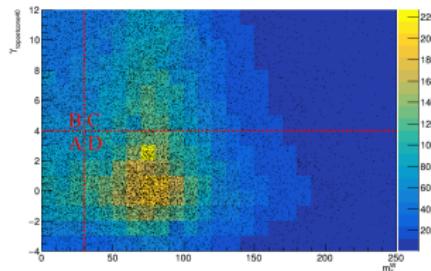
$e$  channel



$\mu$  channel



## ► Unconverted Photons



Channel:	Converted	Unconverted
Electron Channel	$1.28 \pm 0.34$	$1.99 \pm 0.52$
Muon Channel	$1.23 \pm 0.50$	$2.27 \pm 0.92$

# Neural Networks

- ▶ Advanced pattern recognition used to classify events
- ▶ A dense neural network is used with various low and high level variable inputs
- ▶ Supervised learning used to approximate any multidimensional function

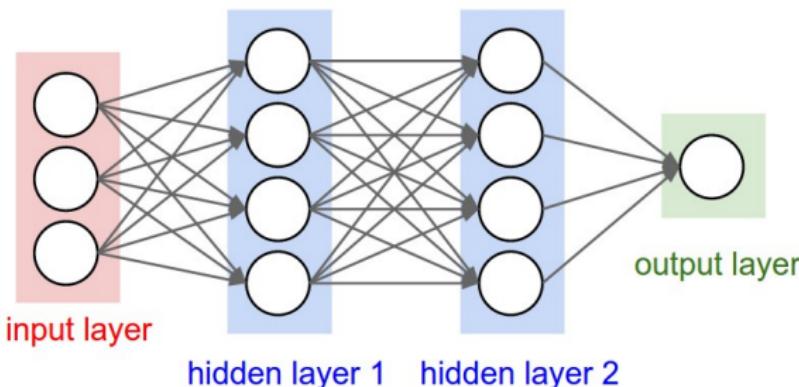


Figure: [Ref: Neural Network]

# Neural Network Optimizers

- ▶ Various optimization functions can be used, I make use of Adam (Adaptive Moment Estimation)
  - ▶ Adam computes adaptive learning rates for every parameter and stores a history of the parameters used to calculate the next step
  - ▶ Stores first (mean) and second (uncentered variance) moments of the gradients used during training
  - ▶ Converges very fast and is less computationally intensive

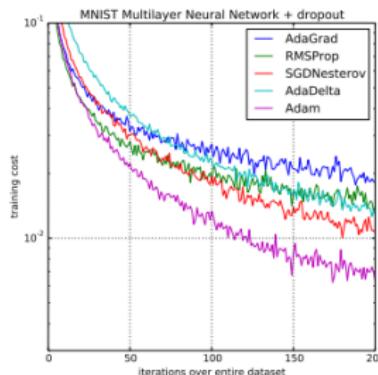


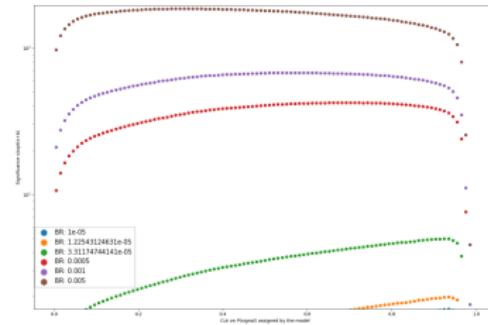
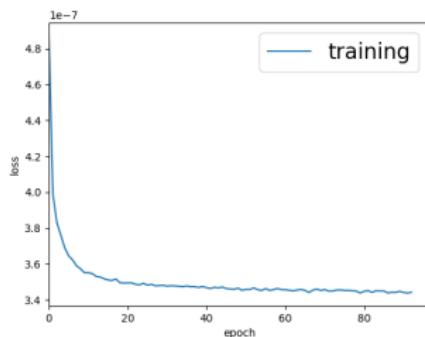
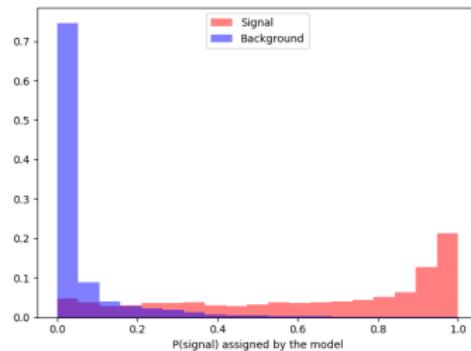
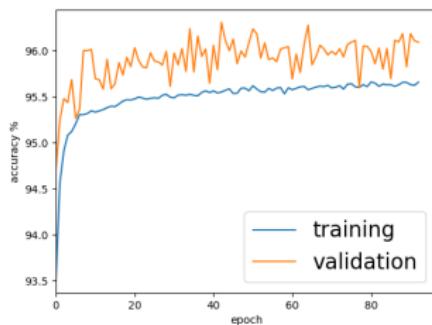
Figure: [Ref:Machine Learning Mastry]

# Neural Network Model Inputs

- ▶ Using keras on top of tensorflow various network architectures were tested
- ▶ Networks are set up with 1 input layer, 2 hidden layers of 10 nodes, and 1 output node
- ▶ Each hidden layer has 20% dropout to prevent overtraining by removing codependency between nodes
- ▶ Batch size of 100 used and each network is allowed 200 epochs (with patience=50), all models converge and end early with reasonable batch sizes
- ▶ Optimizer: Adam
- ▶ Loss Function: Binary Cross Entropy

# Neural Network Example Outputs

e+jets Channel Example



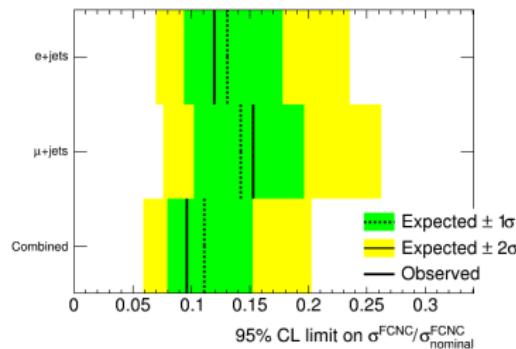
# Analysis

# Signal Region Plots

# Systematic Uncertainties - Nuisance Parameters

# CL<sub>s</sub> Method - Limit Setting

# Result - Limit on $\text{BR}(t \rightarrow q\gamma)$



Channel	Obs. Limit	Exp. Limit
e+jets	$1.19 \times 10^{-4}$	$1.31^{(+0.47)}_{(-0.36)} \times 10^{-4}$
$\mu + \text{jets}$	$1.53 \times 10^{-4}$	$1.42^{(+0.51)}_{(-0.39)} \times 10^{-4}$
Combined	$0.96 \times 10^{-4}$	$1.10^{(+0.43)}_{(-0.30)} \times 10^{-4}$

# Conclusions

- ▶ Limits have been set on the process  $t \rightarrow q\gamma$
- ▶ Neural network implementation improved result by up to 30% compared to stats only improvement

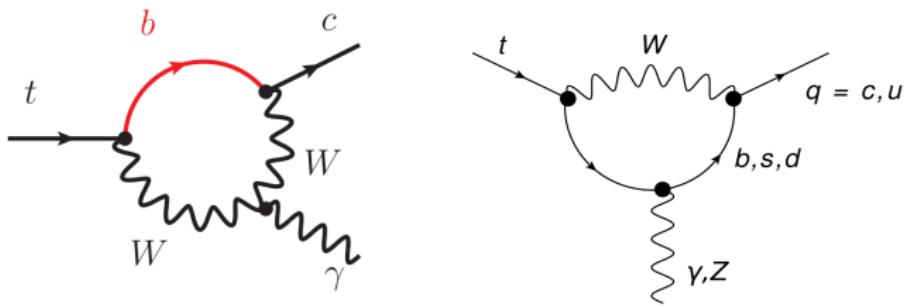
# Thank You

Special thank you to my committee:  
David Strom (Chair)  
Spencer Chang  
Dev Sinha  
My Advisor: Jim Brau

# Questions?

# Backup

# FCNC Diagrams



# Neural Network Model Inputs

$$\text{Separation} = \sum_i^{bins} \frac{n_{si} - n_{bi}}{n_{si} + n_{bi}}$$

## mu+jets channel

Variable	Separation
photon0iso	41.18
mqgam	28.27
photon0pt	24.07
mtSM	11.60
mlgam	7.56
deltaRjgam	5.64
deltaRbl	4.42
MWT	3.34
ST	3.30
nuchi2	3.12
jet0pt	2.81
njets	2.07
smchi2	1.89
wchi2	1.87
jet0e	1.52
deltaRlgam	1.17
leptone	0.87
deltaRjb	0.86
met	0.68
bjet0pt	0.52
leptoniso	0.27

## e+jets channel

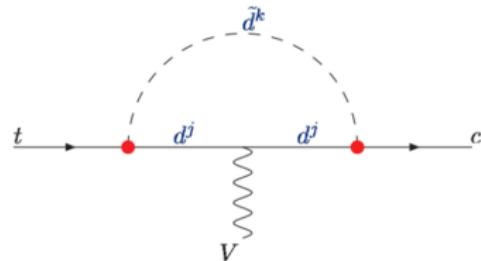
Variable	Separation
photon0pt	23.14
mqgam	22.73
photon0iso	18.70
mtSM	11.02
mlgam	9.53
deltaRbl	5.00
deltaRjgam	4.60
ST	3.83
MWT	3.16
jet0pt	2.47
njets	1.70
nuchi2	1.59
deltaRlgam	1.40
wchi2	1.33
smchi2	1.09
deltaRjb	0.88
leptone	0.85
leptoniso	0.56
bjet0pt	0.50
met	0.47

# Input Variables

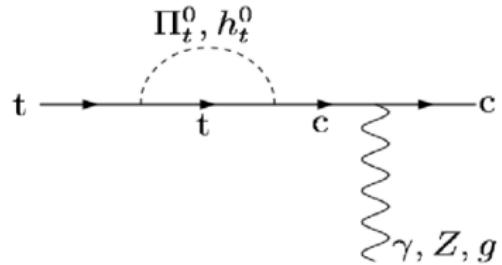
```
['photon0iso','photon0pt','mqgam','mlgam','mtSM','deltaRjgam','deltaRbl','MWT','ST','njets','wchi2','jet0pt','deltaRlgam','leptone','met','bjet0pt']
```

# A Couple BSM Diagrams

- R-parity-violating supersymmetric models  
[\[arXiv:hep-ph/9705341\]](https://arxiv.org/abs/hep-ph/9705341)



- Top-color-assisted technicolor models  
[\[arXiv:hep-ph/0303122\]](https://arxiv.org/abs/hep-ph/0303122)



# Jets/AntiKT

$$d_{ij} = \min\left(\frac{1}{p_{ti}^2}, \frac{1}{p_{tj}^2}\right) \frac{\Delta_{ij}^2}{R^2}$$

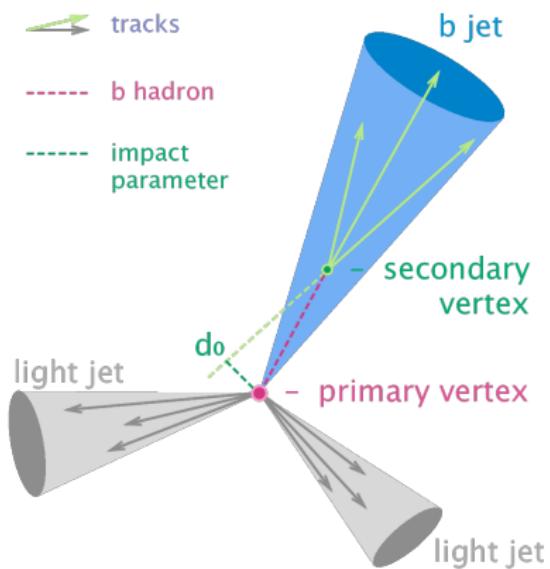
$$d_{iB} = \frac{1}{p_{ti}^2}$$

$$\Delta_{ij}^2 = (\eta_i - \eta_j)^2 + (\phi_i - \phi_j)^2$$

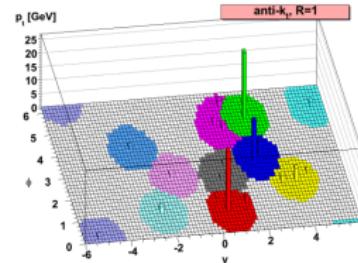
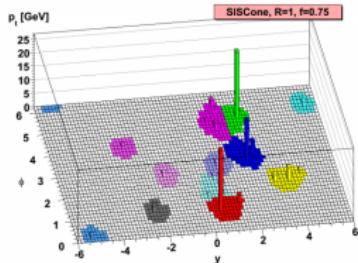
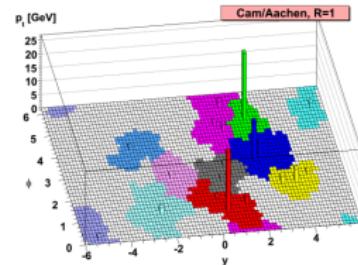
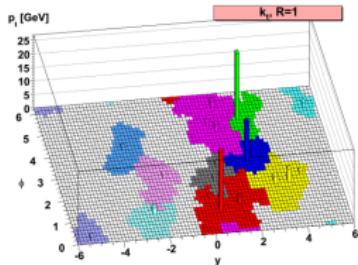
- ▶ Find minimum of entire set of  $\{d_{ij}, d_{iB}\}$
- ▶ If  $d_{ij}$  is the minimum particles i,j are combined into one particle and removed from the list of particles
- ▶ If  $d_{iB}$  is the minimum i is labelled as a final jet and removed from the list of particles
- ▶ Repeat until all particles are part of a jet with distance between jet axes  $\Delta_{ij}$  is greater than R

# B-tagging

- ▶ B Hadrons travel a measurable distance before decay
- ▶ Tracks originate from outside of interaction point (Secondary Vertex)
- ▶ Backtracking tracks in displaced vertex gives an impact parameter
- ▶ Decay chain MVA attempts to reconstruct decay of the jet
- ▶ Outputs of these algorithms used in a BDT to determine if a Jet is from a b-quark



# Jet Algorithms



- ▶ IR Safety: Adding soft emission particle does not change final jet configuration
- ▶ Collinear Safety: Splitting a jet into 2 collinear jets yields the same result, jet size does not matter

# Jet Algorithms

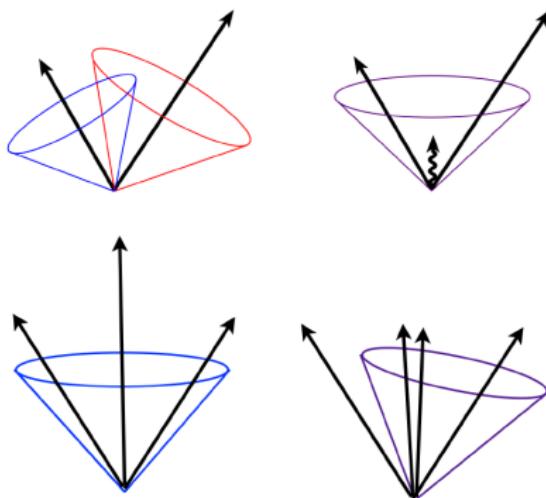
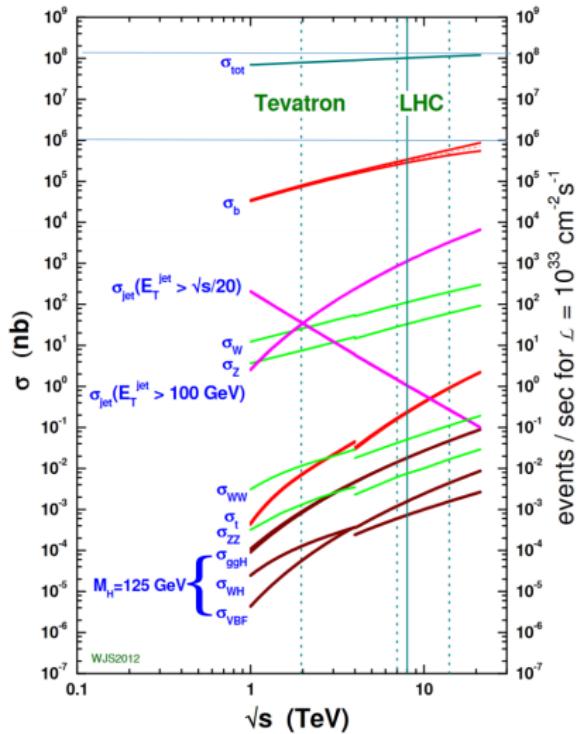


FIGURE 4.4. Illustration of the infrared sensitivity of a cursory designed jet algorithm (top). Illustration of the product of a collinear unsafe jet algorithm. A collinear splitting changes the number of jets (bottom). [52].

Courtesy - John Myers

# What's in the data?

proton - (anti)proton cross sections

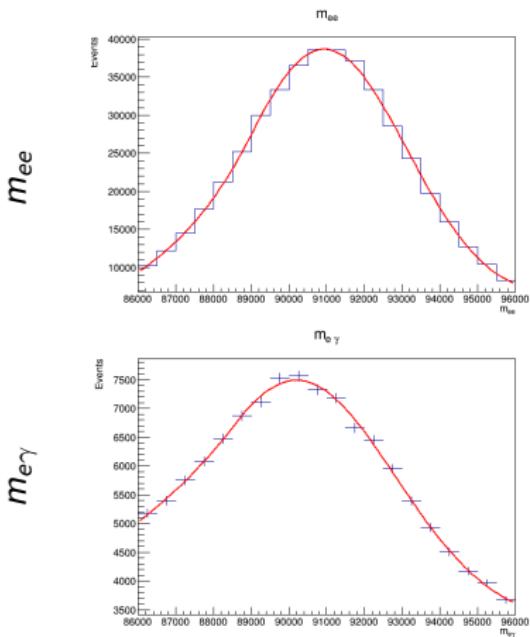


- ▶ The number of events we see is  $N = \sigma L$
- ▶  $\sigma_{t\bar{t}} = 831.76 \text{ pb}$
- ▶  $N_{t\bar{t}} \approx 116 \times 10^6$
- ▶  $N_{\text{tot}} \approx 14 \times 10^{15}$  events produced during the 13 TeV data runs

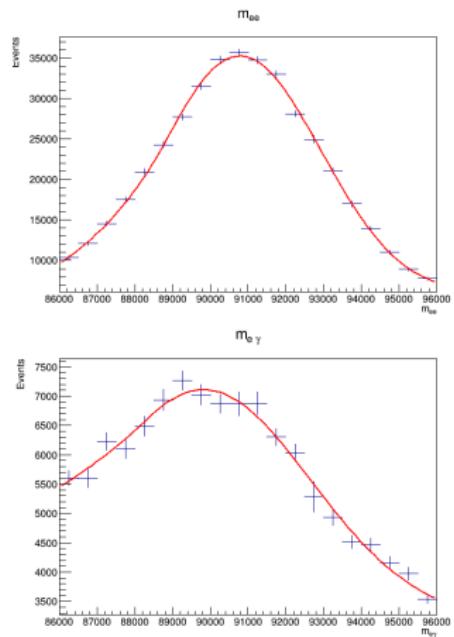
$m_{ee}, m_{e\gamma}$ 

## egamma fake rate Data and MC

## ► Data

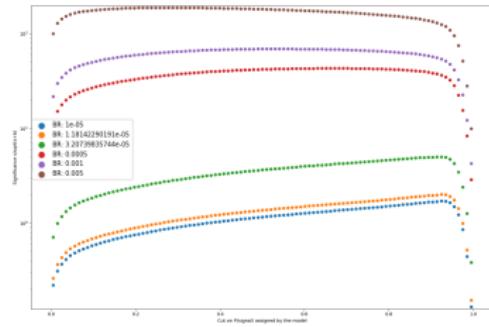
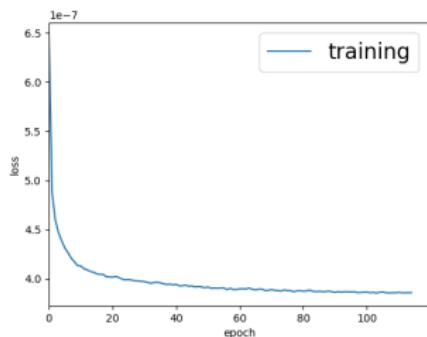
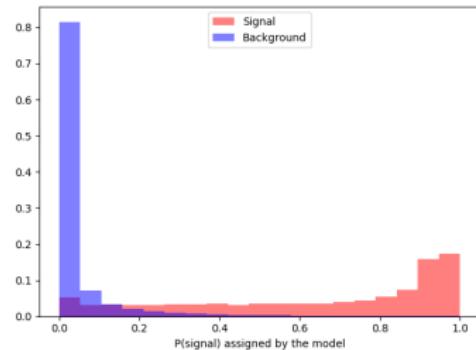
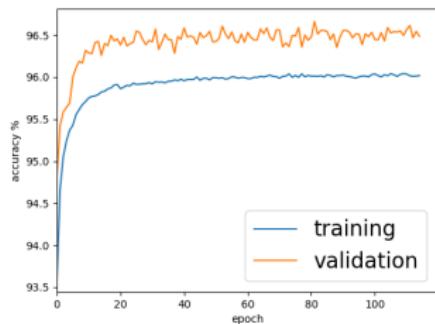


## ► Monte Carlo



# Neural Network Example Outputs

## mu+jets Channel Example



# Duplicate Event Removal

$$\mathcal{L}_{tq\gamma}^{eff} = -e\bar{c}\frac{i\sigma^{\mu\nu}q_\nu}{m_t}(\lambda_{ct}^L P_L + \lambda_{ct}^R P_R)tA_\mu + H.c.$$