

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

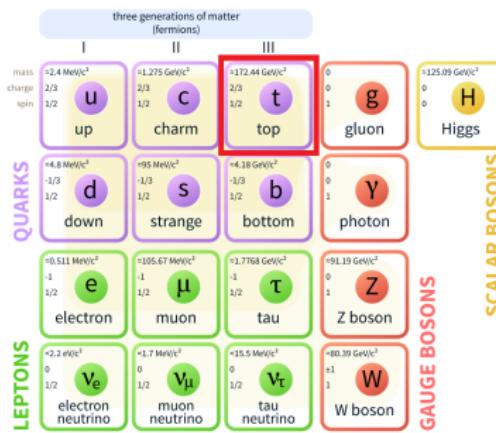
The LHC and ATLAS

Searching for Flavor-Changing Neutral Current Signatures

Conclusion

The Standard Model

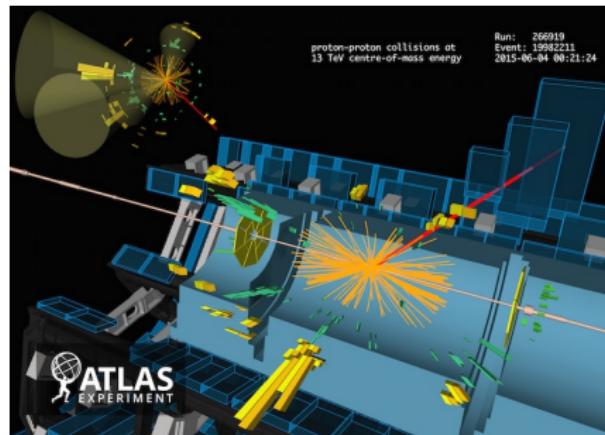
Standard Model of Elementary Particles



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

The Top Quark

- ▶ Heaviest fundamental particle, 172.5 GeV
- ▶ Lifetime 5×10^{-25} 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 at the LHC
 - ▶ Quark-antiquark annihilation $\approx 10\%$
 - ▶ Gluon-gluon fusion $\approx 90\%$

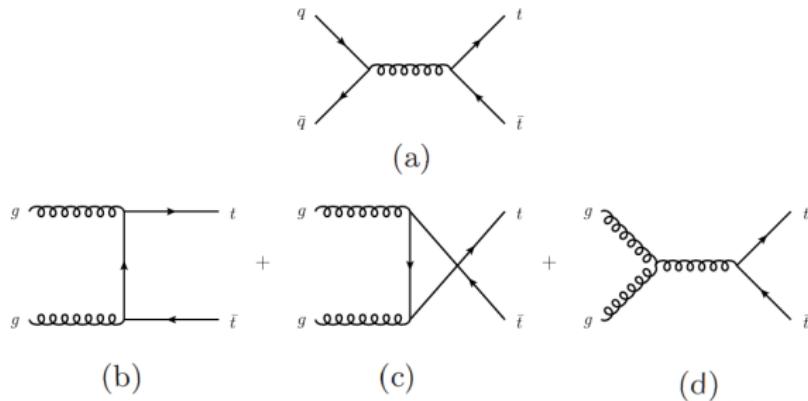


Figure: Leading order $t\bar{t}$ diagrams

Top Quark Pair Production

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

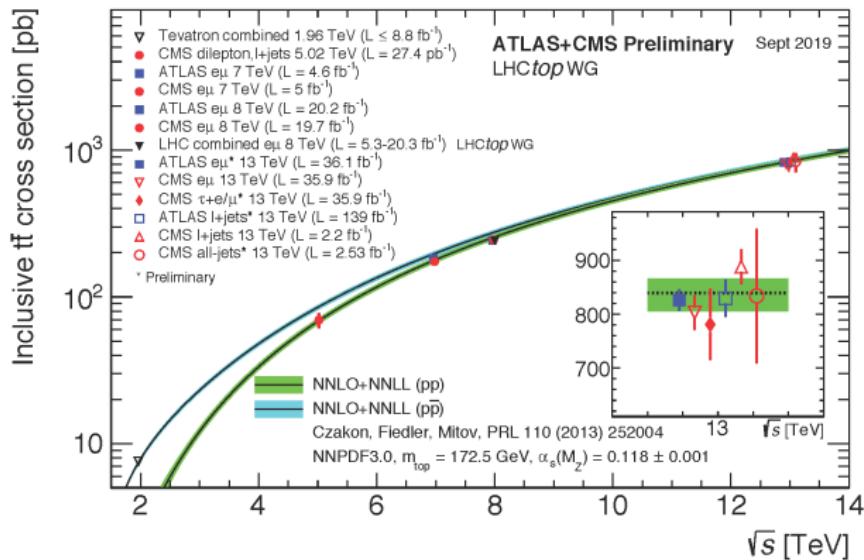


Figure: $t\bar{t}$ production cross section [TopWGSummaryPlots]

Top Quark Decays

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

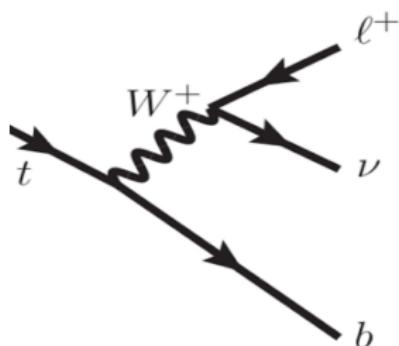


Figure: Leptonic final state diagram for a top decay

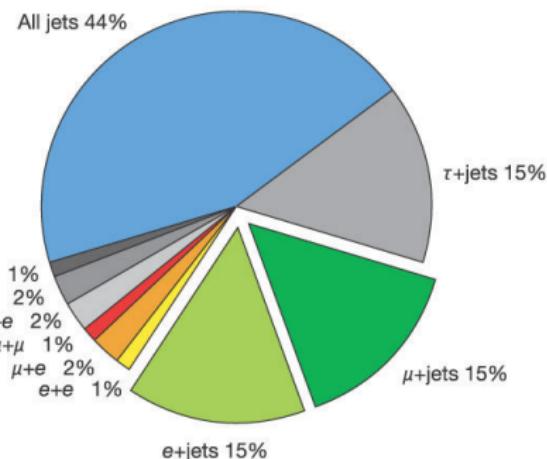
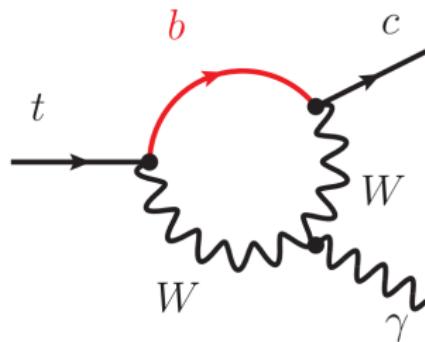
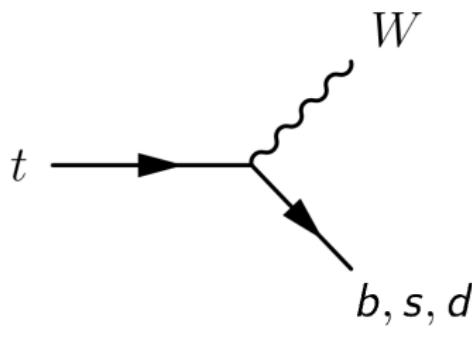


Figure: Top quark pair decay final states [Nature]

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 $\text{BR}(t \rightarrow \gamma q)$ processes: [Phys.Lett. B800 135082]
 - ▶ $t \rightarrow \gamma u < 2.8 \times 10^{-5}$
 - ▶ $t \rightarrow \gamma c < 1.8 \times 10^{-4}$

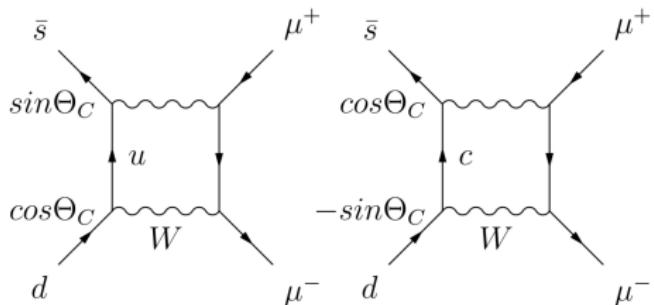
GIM Mechanism

- ▶ 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

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

- ▶ 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 currents/natural suppression of neutral currents

GIM Mechanism



- ▶ 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

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

- ▶ Decay rates proportional to $|V_{tx}|^2$
- ▶ 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

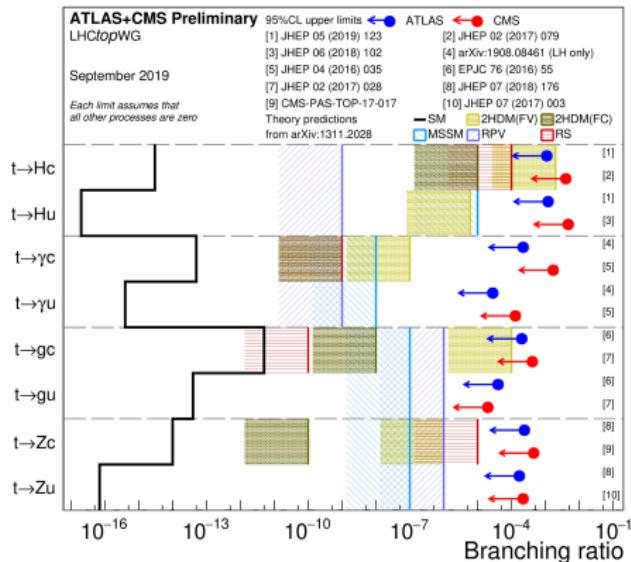
Top Flavor Changing Neutral Currents (FCNCs)

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

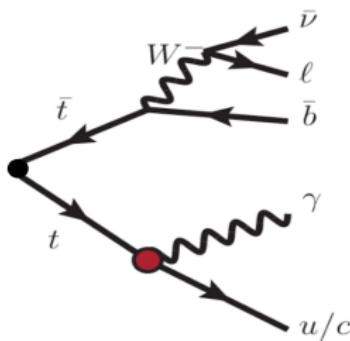
- $t \rightarrow \gamma u < 2.8 \times 10^{-5}$
- $t \rightarrow \gamma c < 18 \times 10^{-5}$

Monte Carlo Production of FCNC Signal Samples

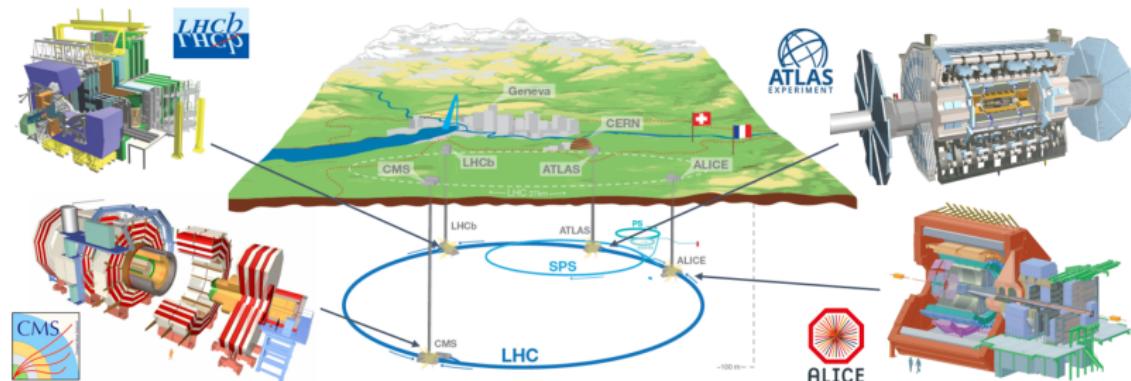
- ▶ Due to the low cross sections must create my own Monte Carlo Samples for the Signal
- ▶ An effective field theory approach was taken in the creation of the model [Degrande et al. Phys. Rev. D 91, 034024 (2015)]
- ▶ This model takes advantage of dimension-6 operators

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

- ▶ Final state topology
 - ▶ One Neutrino, from W
 - ▶ One Lepton (e or μ), from W
 - ▶ One B-jet, SM Top
 - ▶ One Photon, FCNC Top
 - ▶ One Jet, FCNC Top

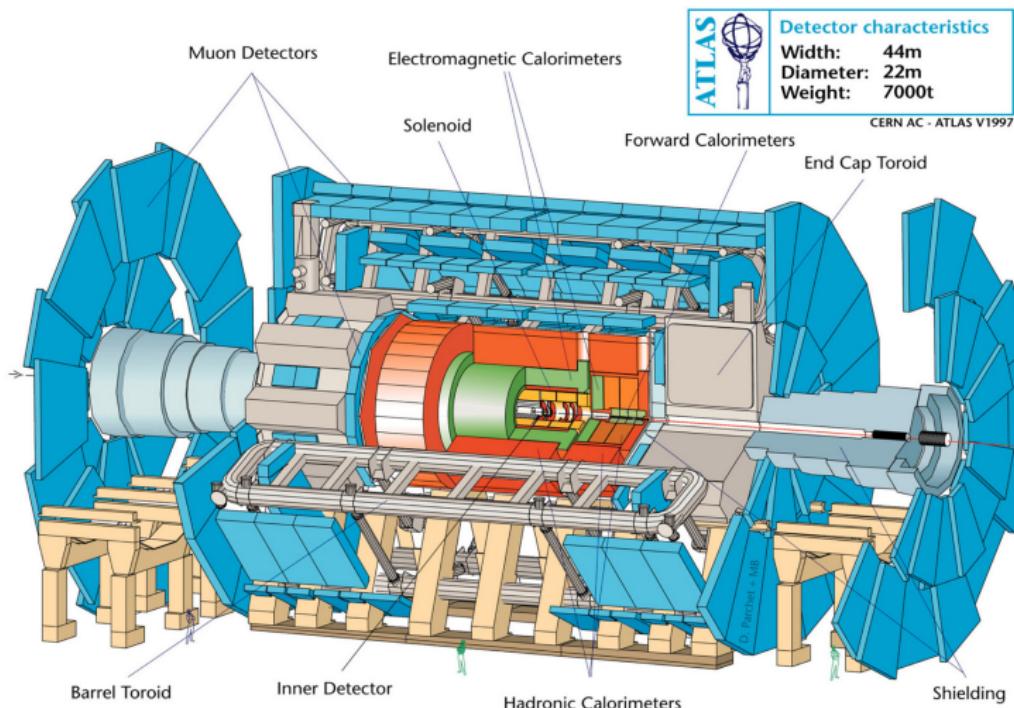


The Large Hadron Collider

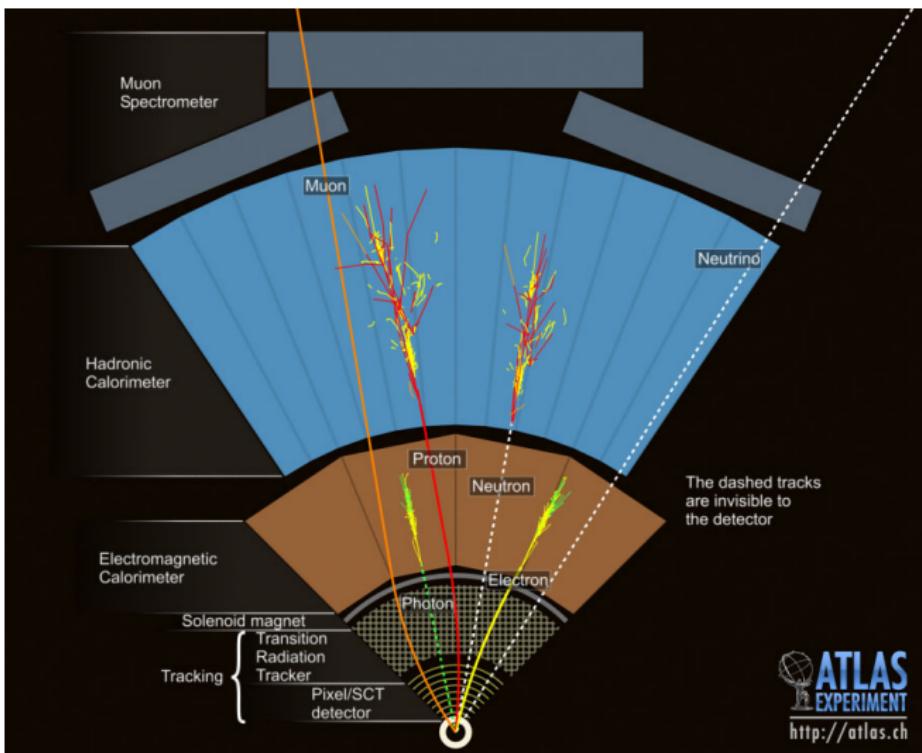


- ▶ 27km ring beneath Franco-Swiss border
- ▶ Collides protons at center of mass energy 13TeV

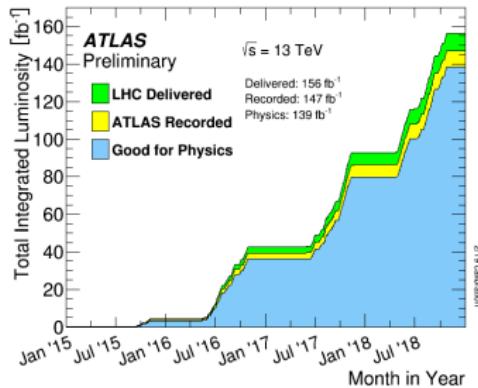
The ATLAS Detector



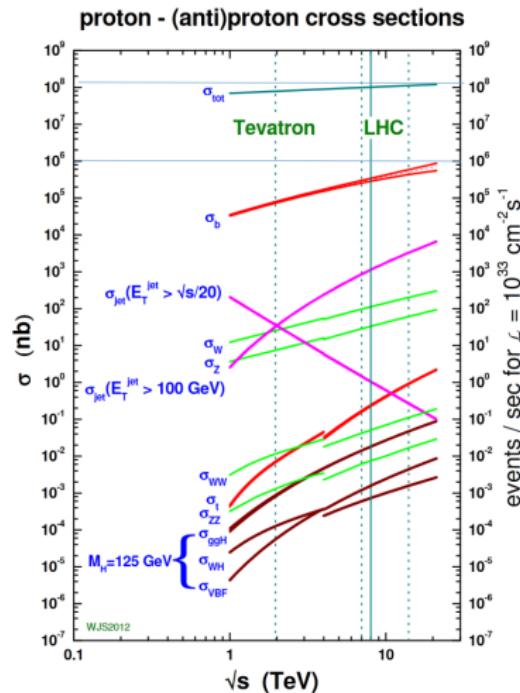
Particles in ATLAS



ATLAS Data

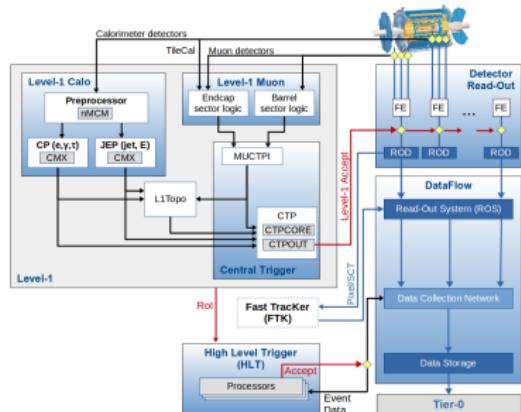
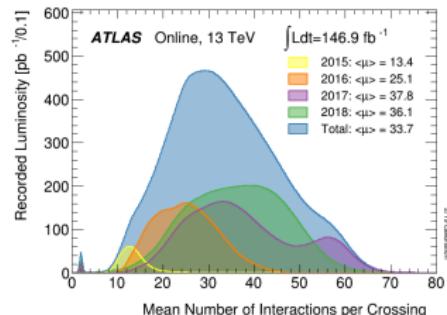


- ▶ Total Luminosity at $\sqrt{s} = 13$ TeV:
 139fb^{-1}
- ▶ The number of events we see is
 $N = \sigma L$
- ▶ $N_{tot} \approx 14 \times 10^{15}$ events produced during the 13TeV data runs
- ▶ $N_{t\bar{t}} \approx 116 \times 10^6$



ATLAS Data Acquisition

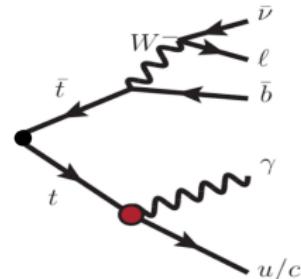
- ▶ Average 33.7 collisions per crossing during Run 2, 40MHz collision rate
- ▶ Raw uncompressed event: 1.6 MB
- ▶ Around 64 TB/s data
- ▶ 2 Stage Trigger System
 - ▶ Level 1 - Hardware based, coarse object reconstruction, reduces rate to under 100 kHz
 - ▶ High-Level Trigger: Software based, performs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



Signal Region Preselection

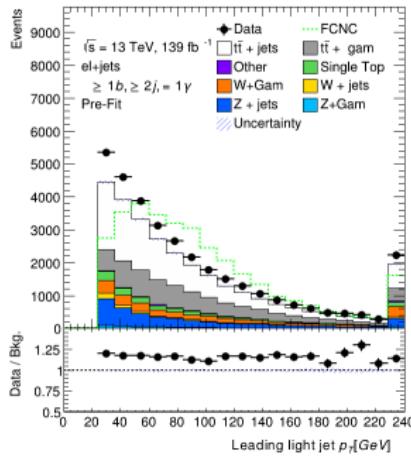
Require:

- ▶ Exactly one lepton (e or μ) ≥ 25 GeV
- ▶ Exactly one photon ≥ 50 GeV
- ▶ Missing transverse energy ≥ 30 GeV
- ▶ ≥ 2 jets (at least 1 b-tag)
 - ▶ Due to all of the processes at hadron colliders it is important to model similar event topologies well.
 - ▶ Major backgrounds: $t\bar{t}$, W+Jets, and processes with an associated photon
 - ▶ Minor backgrounds: Z+Jets, Single Top

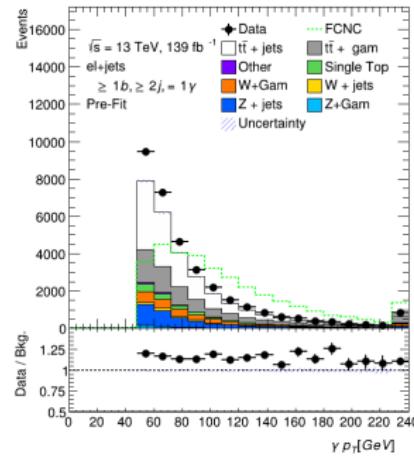


Preselection Objects, e+jets channel

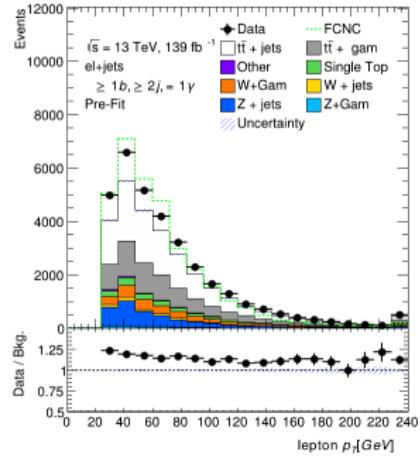
► Leading Jet p_T



► Photon p_T

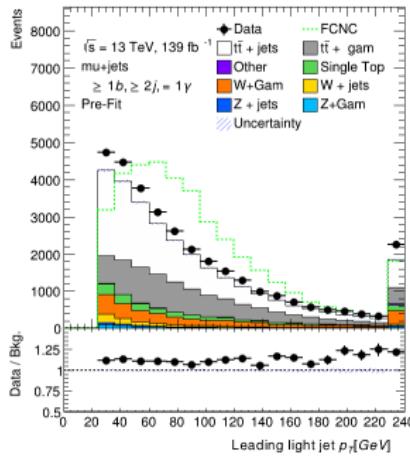


► Lepton p_T

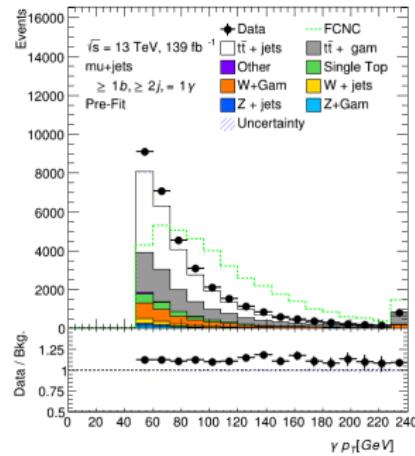


Preselection Objects, $\mu + \text{jets}$ channel

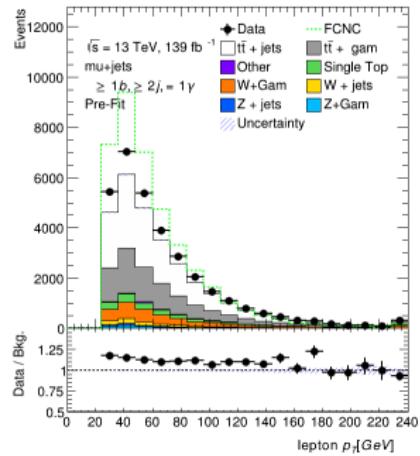
► Leading Jet p_T



► Photon p_T



► Lepton p_T

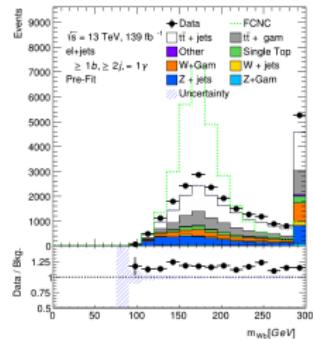


Top Quarks

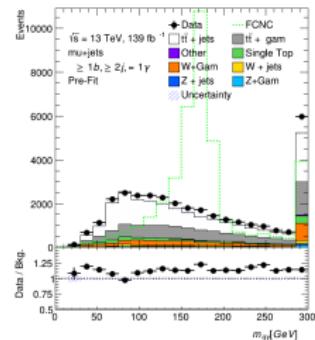
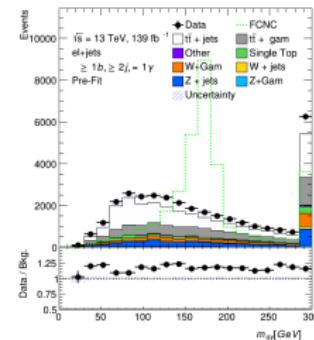
Tops must be reconstructed from these basic physics objects along with b-jets and E_T^{miss}

- Neutrino z-axis direction is ambiguous, determined by minimization of: $\chi^2 = \frac{(m_{b,l,\nu} - m_t)^2}{\sigma_{SMtop}^2} + \frac{(m_{l,\nu} - m_W)^2}{\sigma_W^2}$

SM Top: m_{Wb}



FCNC Top: $m_{q\gamma}$



Additional Processes

- ▶ QCD Backgrounds are difficult to model at the energies and interaction rates of the LHC
 - ▶ Develop 0 photon regions for scale factors for major backgrounds ($t\bar{t}$ and W+jets)
- ▶ Photon misidentification i.e., electrons or hadrons being reconstructed as photons ($e \rightarrow \gamma$, $j \rightarrow \gamma$)
 - ▶ $e \rightarrow \gamma$: $Z \rightarrow ee$ tag and probe method used for SF calculation
 - ▶ $j \rightarrow \gamma$: ABCD method used for SF calculation

0 Photon Backgrounds

Select events with objects that look like similar to our expected topology for major backgrounds ($t\bar{t}$, W+jets) Require:

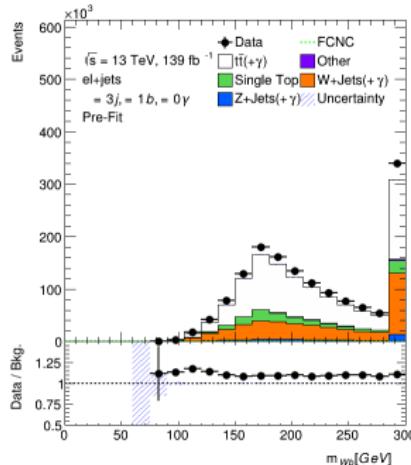
- ▶ Common Object Selection (MET, Triggers, etc.)
- ▶ Exactly one lepton (e or μ) ≥ 25 GeV
- ▶ Number of Jets to define Regions:
 - ▶ W+jets enriched region: $n_{\text{jets}} = 3$
 - ▶ Validation Region: $n_{\text{jets}} = 4$
 - ▶ $t\bar{t}$ +jets enriched region: $n_{\text{jets}} \geq 5$
- ▶ Exactly 1 b-tagged jet

Calculate scale factors simultaneously:

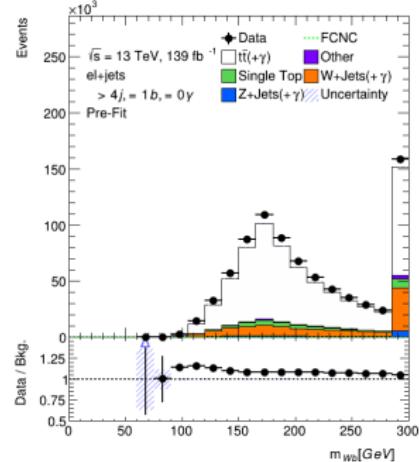
$$\begin{bmatrix} N(W)_{3j} & N(t\bar{t})_{3j} \\ N(W)_{5+j} & N(t\bar{t})_{5+j} \end{bmatrix} \begin{bmatrix} W_{SF} \\ t\bar{t}_{SF} \end{bmatrix} = \begin{bmatrix} N(\text{data-bkg})_{3j} \\ N(\text{data-bkg})_{5j} \end{bmatrix}$$

0 Photon SFs, m_{Wb} before SF, electron channel

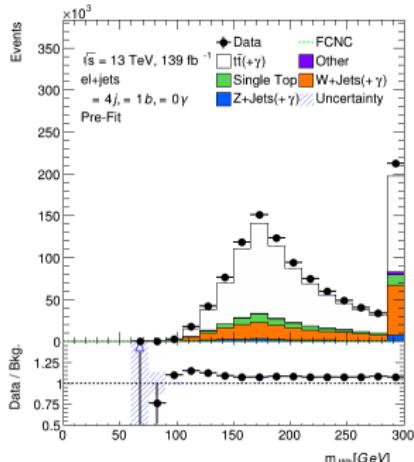
► W+jets Enriched



► $t\bar{t}$ Enriched



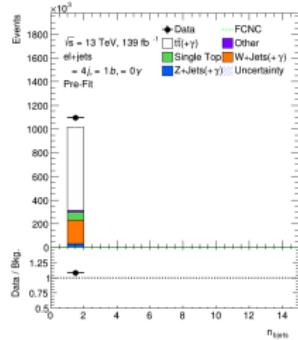
► Validation Region



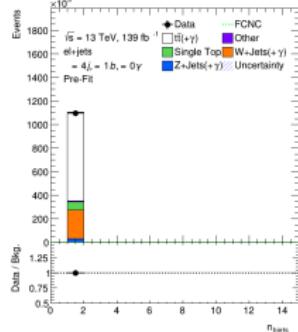
Sample	e+jets SF	μ +jets SF
W+jets	1.22	1.25
$t\bar{t}$	1.06	1.01

0 Photon Regions: VR before/after SF, electron channel

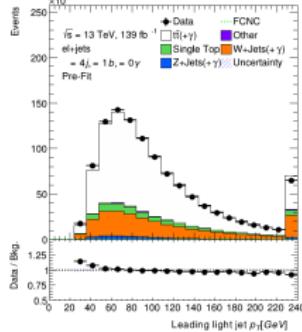
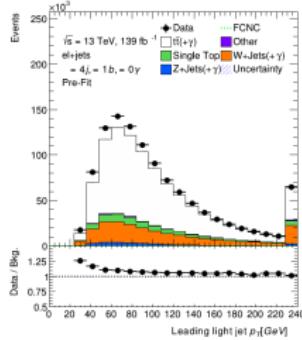
► n_{bjets}



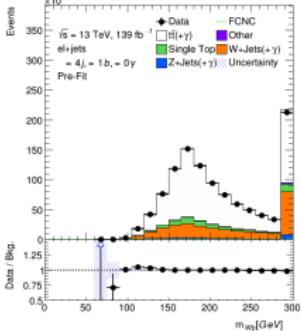
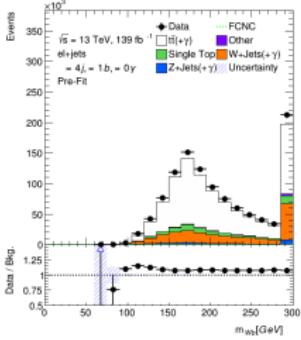
After SF



► Lead Jet p_T



► m_{Wb}



$j \rightarrow \gamma$ Fake Rate Scale Factor: ABCD Method

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.

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

!Isolated



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

Isolated

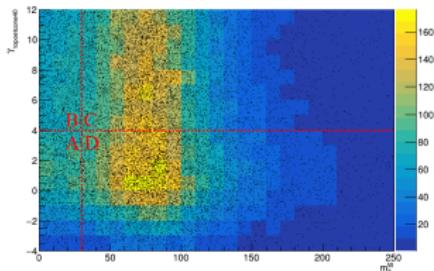


$$\text{SF}^{\text{h-fake}} = \frac{N_{D,\text{est.}}^{\text{h-fake}}}{N_{D,\text{MC}}^{\text{h-fake}}}$$

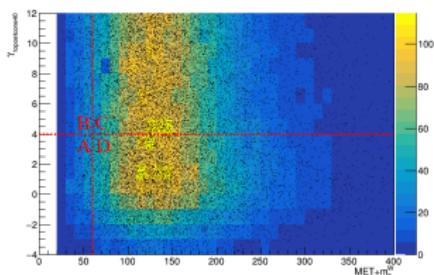
Reversed MET,MWT SR MET,MWT

► Converted Photons

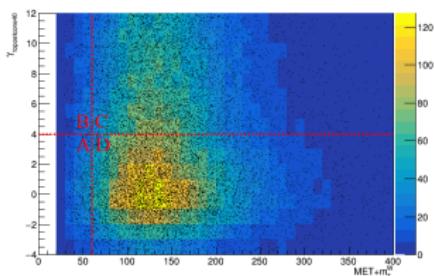
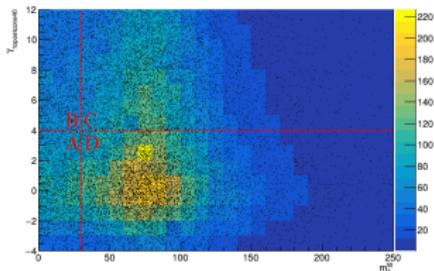
e channel



μ channel



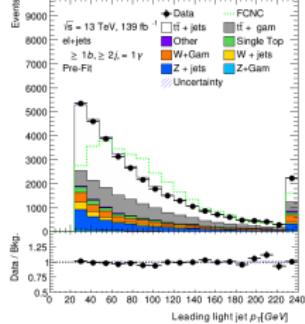
► Unconverted Photons



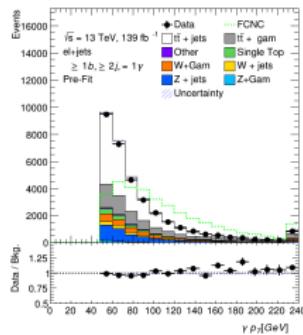
Channel:	Converted	Unconverted
Electron Channel	1.28 ± 0.34	1.99 ± 0.52
Muon Channel	1.23 ± 0.50	2.27 ± 0.92

Post SF Preselection Objects

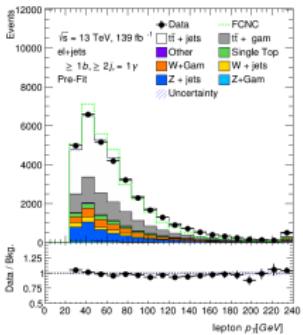
Electron Channel



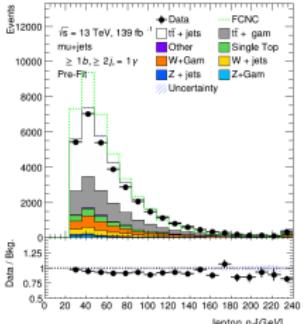
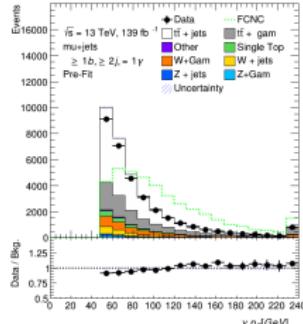
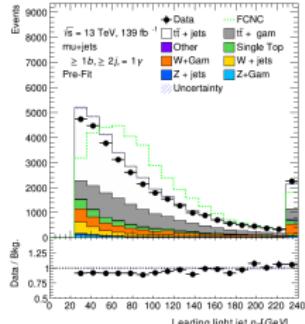
Photon p_T



Lepton p_T



Muon Channel



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

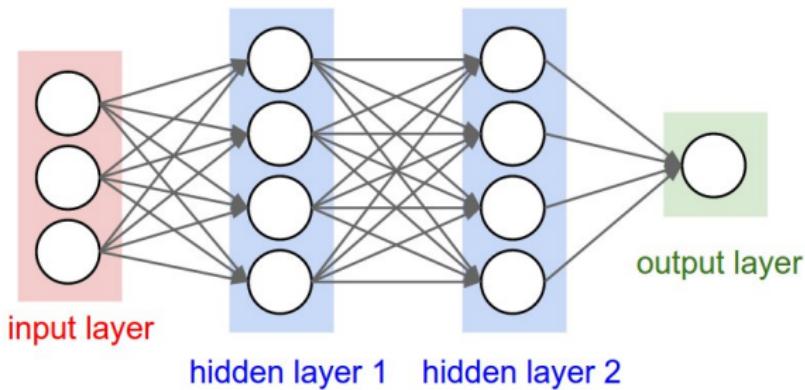
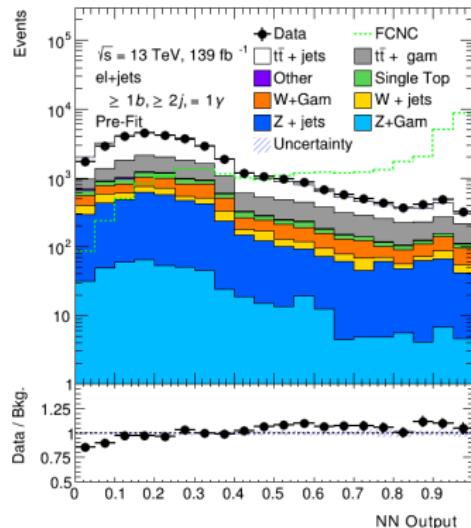


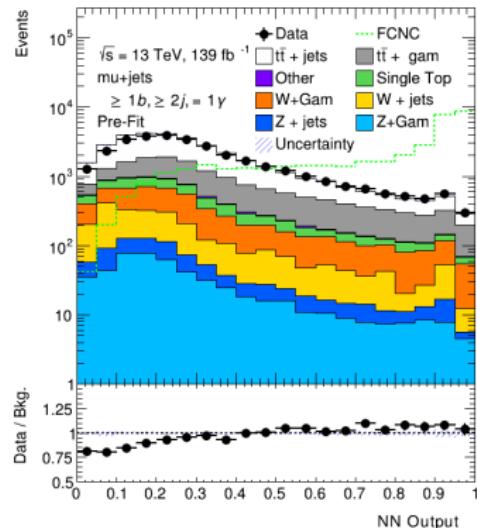
Figure: [Ref: Neural Network]

Scaled Preselection Neural Network Outputs

Electron Channel



Muon Channel



Deviations not unexpected, all SFs optimized for SR

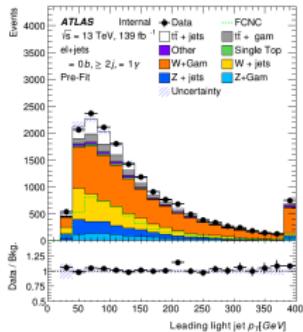
Signal And Validation Regions

Validate the SFs in regions closer to the Signal Region

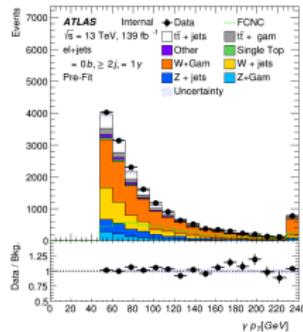
Region:	SR	W+ γ VR	$t\bar{t} + \gamma$ VR
$n_\gamma \geq 50$ GeV	=1	=1	=1
$n_{\text{lep}} \geq 25$ GeV	=1	=1	=1
$n_{\text{jets}} \geq 25$ GeV	≥ 2	≥ 2	≥ 2
$n_{\text{bjets}} \geq 25$ GeV	=1	=0	≥ 1
NN Cut	\geq	-	<

Validation Regions: $W + \text{jets} + \gamma$

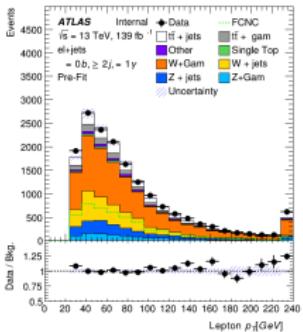
Electron Channel



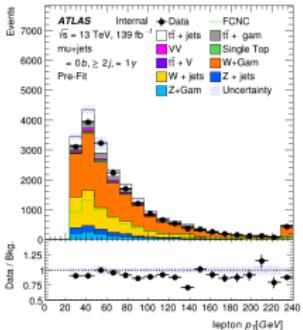
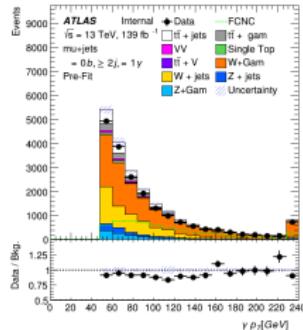
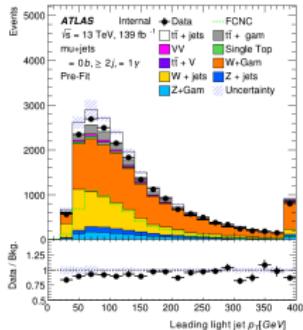
► Photon p_T



► Lepton p_T

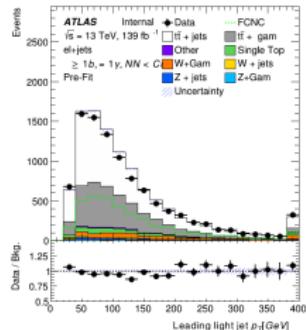


Muon Channel

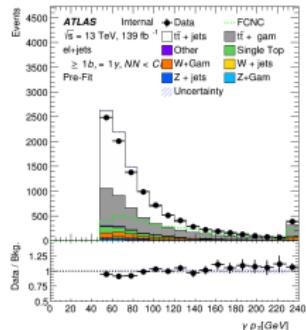


Validation Regions: $t\bar{t} + \text{jets} + \gamma$

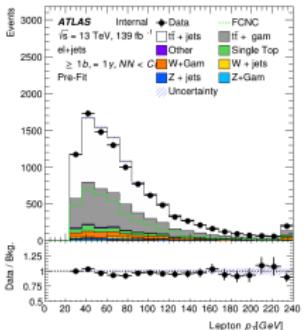
Electron Channel



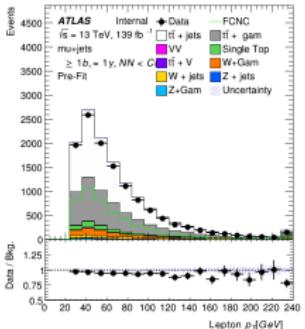
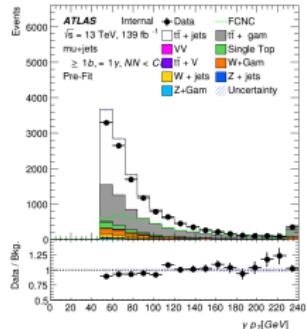
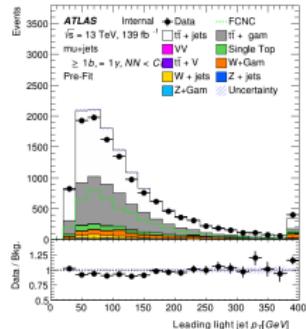
► Photon p_T



► Lepton p_T

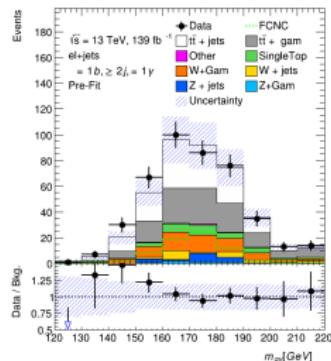


Muon Channel

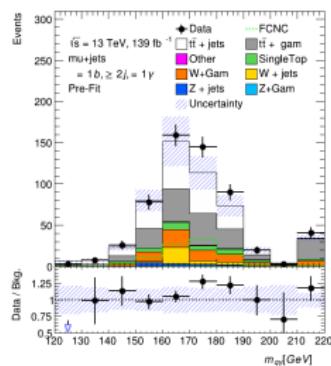


Pre-fit Signal Region Plots

Electron channel



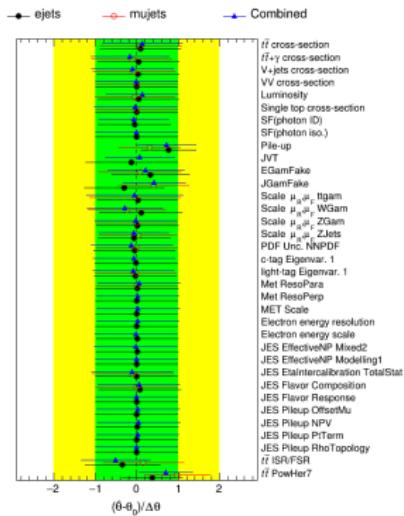
Muon channel



Systematic Uncertainties - Nuisance Parameters

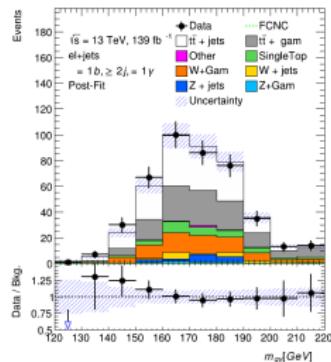
- ▶ Most systematics do not effect fit
- ▶ Largest pulls from Photon and Jet related variables
- ▶ ISR/FSR Variation
- ▶ $j \rightarrow \gamma$ fake rate
- ▶ Changing MC generator

Statistically dominated limit

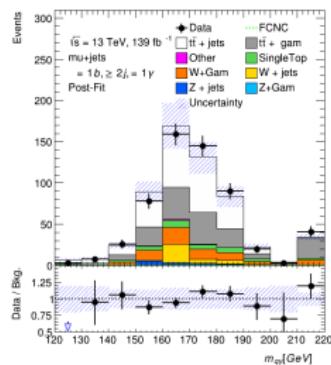


Post-fit Signal Region Plots

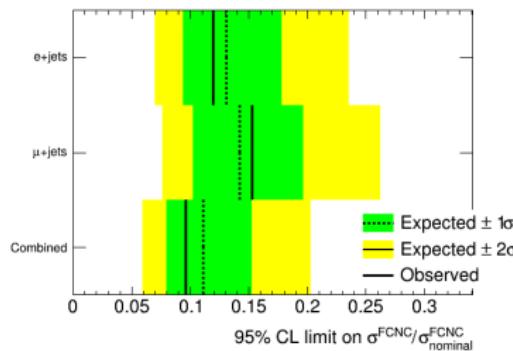
Electron channel



Muon channel



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



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

Conclusions

- ▶ Limits have been set on $\text{BR}(t \rightarrow q\gamma)$, the best current limit on $\text{BR}(t \rightarrow c\gamma) < 9.6 \times 10^{-5}$
- ▶ Previous result in 8TeV Data: $\text{BR}(t \rightarrow q\gamma) < 6.3 \times 10^{-4}$
 - ▶ Expected statistics only improvement: 1.3×10^{-4}
 - ▶ Neural network implementation improved result by up to 30%
- ▶ Continuing to prod the Standard Model from all angles is important in searching for new physics

Thank You

Special thank you to my committee:

David Strom (Chair)

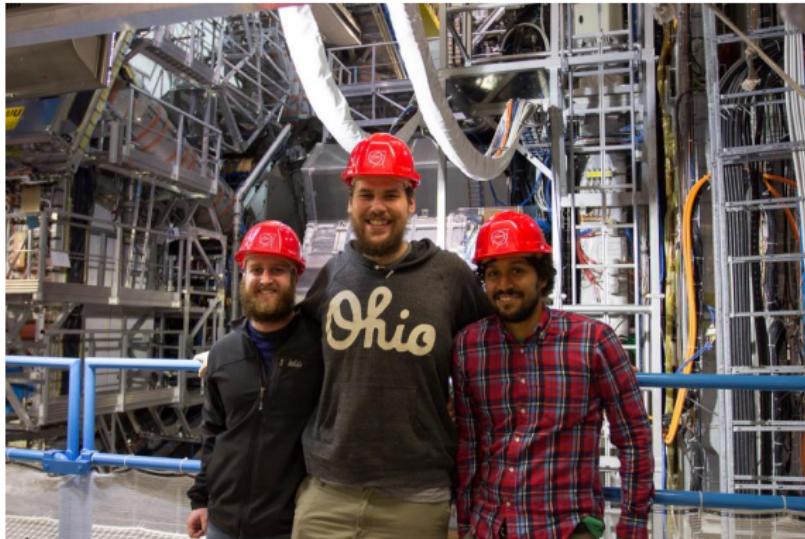
Spencer Chang

Dev Sinha

My Advisor: Jim Brau

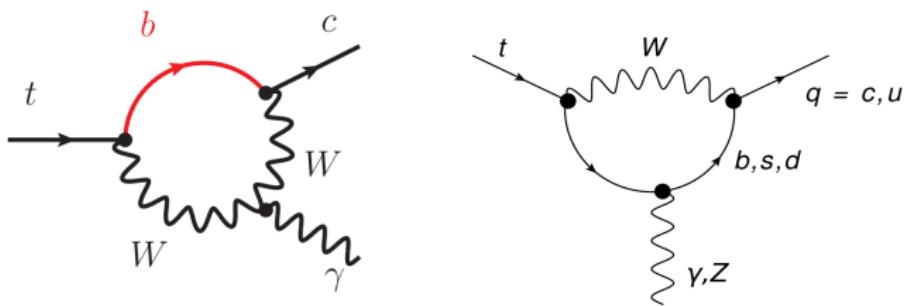
And thank you all for coming!

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
mqgam	28.27
photon0pt	24.07
photon0iso	21.18
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 and Loss Function

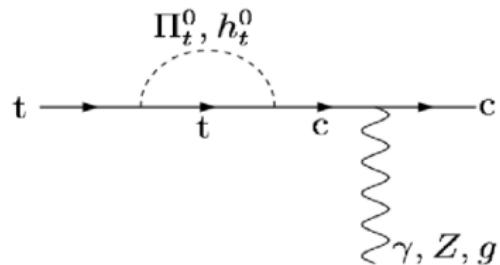
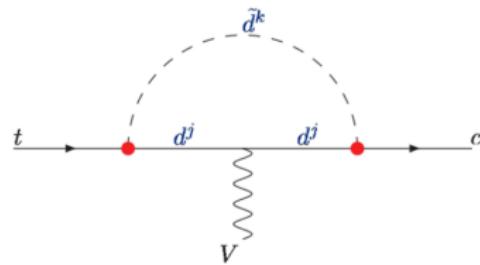
```
['photon0iso','photon0pt','mqgamm','mlgamm','mtSM','deltaRjgamm','deltaRbl','MWT','ST','njets','wchi2','jet0pt','deltaRlgam','leptone','met','bjet0pt']
```

$$\text{Loss} = -\frac{1}{N} \sum_{i=1}^N y_i \log(p(y_i)) + (1 - y_i) \log(1 - p(y_i))$$

- ▶ y - binary indicator (0 or 1) if class label is the correct classification for observation
- ▶ p - predicted probability observation is the class label (0 or 1)

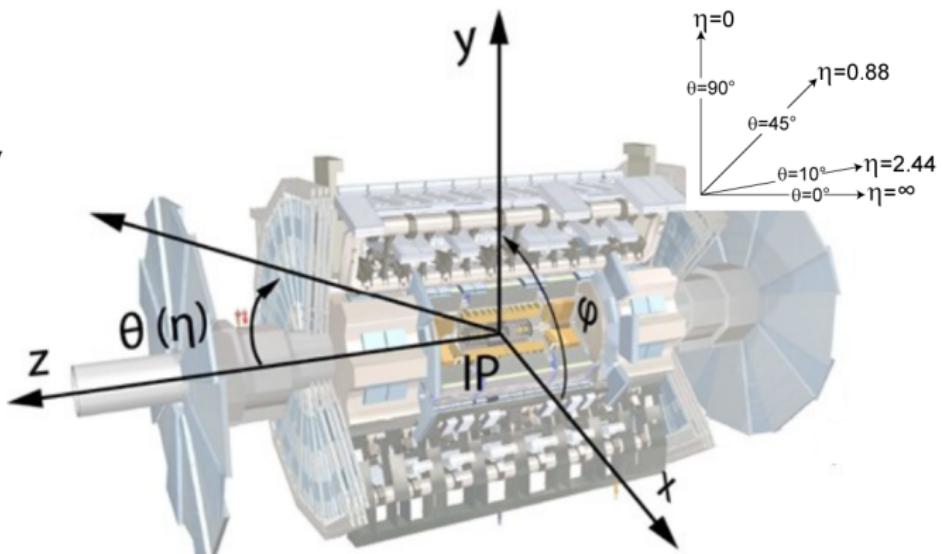
A Couple BSM Diagrams

- R-parity-violating supersymmetric models
[arXiv:hep-ph/9705341]
- Top-color-assisted technicolor models
[arXiv:hep-ph/0303122]



The ATLAS Detector - Coordinates

pseudorapidity
$$\eta = -\ln \tan \frac{\theta}{2}$$



Jets/AntiKT

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

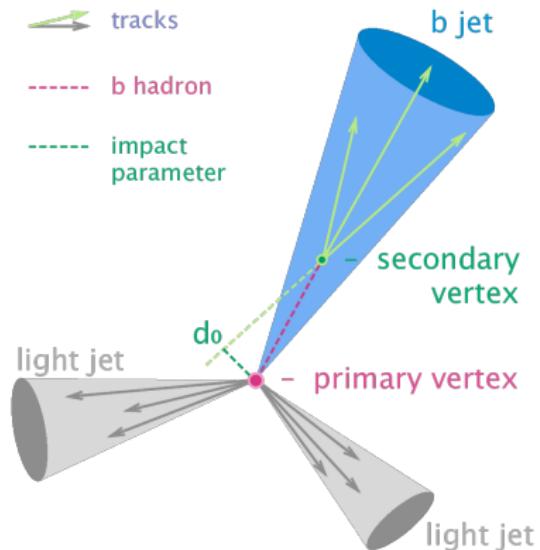
$$d_{iB} = \frac{1}{p_{ti}^2}$$

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

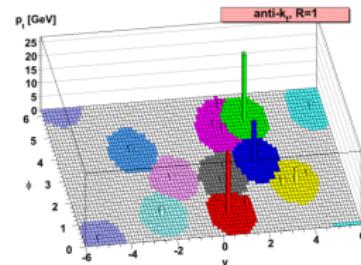
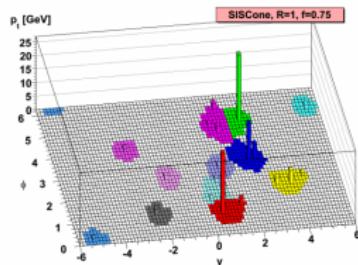
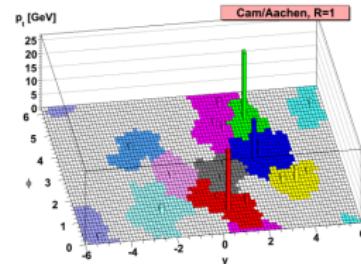
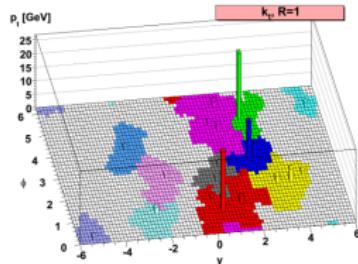
- ▶ Find minimum of entire set of $\{d_{ij}, d_{iB}\}$
- ▶ If d_{ij} is the minimum particles i,j are combined into one particle and removed from the list of particles
- ▶ If d_{iB} is the minimum i is labelled as a final jet and removed from the list of particles
- ▶ Repeat until all particles are part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

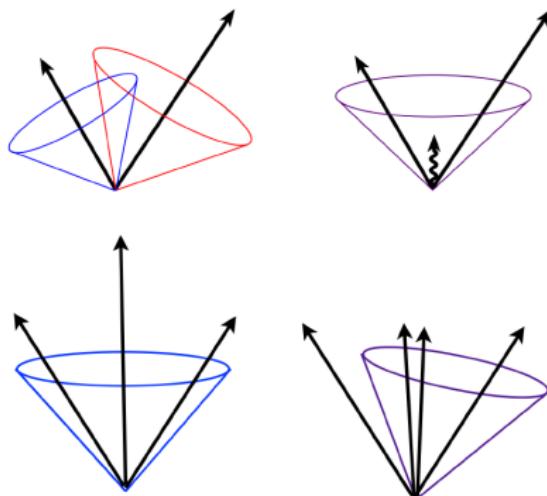
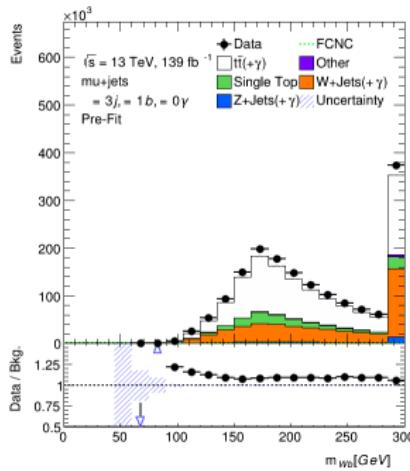


FIGURE 4.4. Illustration of the infrared sensitivity of a cursory designed jet algorithm (top). Illustration of the product of a collinear unsafe jet algorithm. A collinear splitting changes the number of jets (bottom). [52].

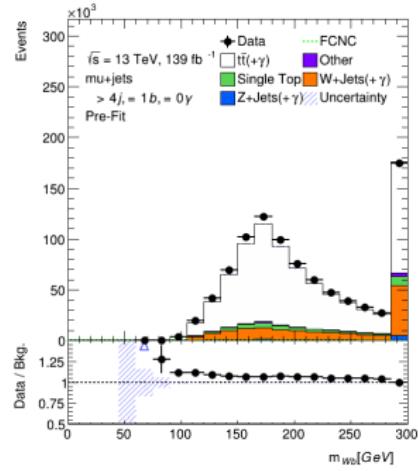
Courtesy - John Myers

0 Photon SFs, m_{Wb} before SF $\mu+jets$ channel

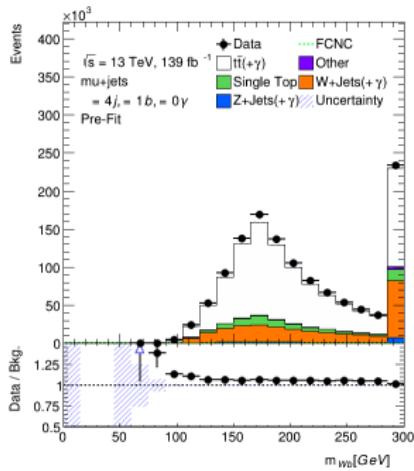
► W+jets Enriched



► $t\bar{t}$ Enriched



► Validation Region



$e \rightarrow \gamma$ 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, ≥ 1 Photon, $86.1 \text{ GeV} < m_{e\gamma} < 96.1 \text{ GeV}$
- ▶ Tag and Probe Method used

$e \rightarrow \gamma$ Scale Factor

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

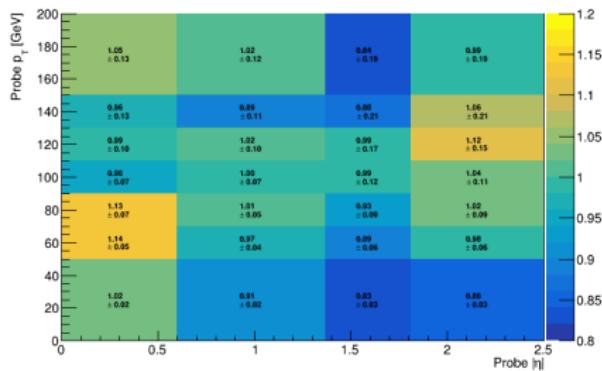
$$\text{SF}_{\text{FR}}^{\text{e-fake}} = \frac{\text{FR}_{\text{data}}^{\text{e-fake}}}{\text{FR}_{\text{MC}}^{\text{e-fake}}}$$

This scale factor is calculated for converted and unconverted photons as well as in bins of η and ϕ

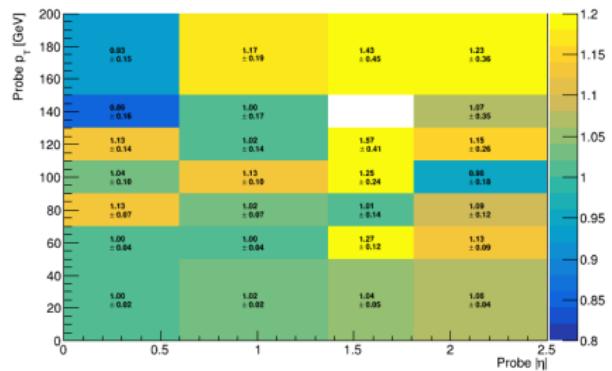
- ▶ Converted photons pair produce before the ECAL leaving tracks in the Inner Detector
- ▶ Unconverted photons only pair produce inside of the ECAL

$e \rightarrow \gamma$ 2D Fake Rates

► Converted γ



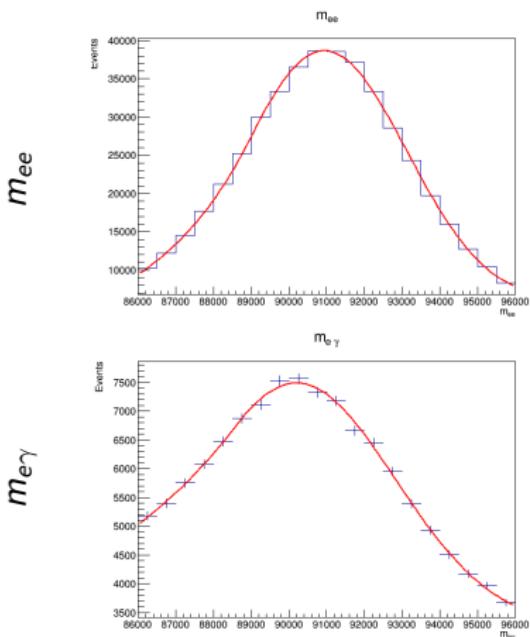
► Unconverted γ



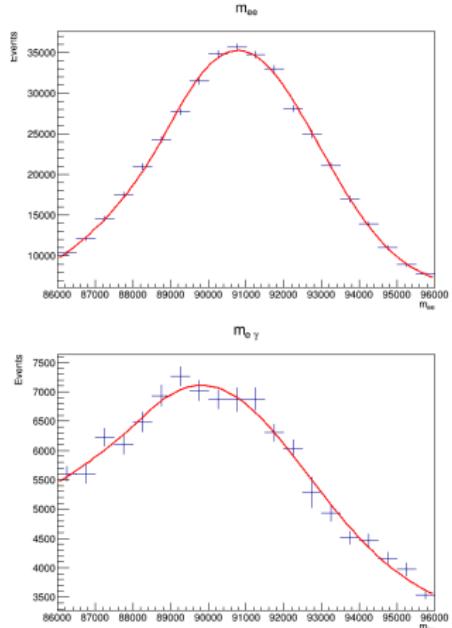
m_{ee} , $m_{e\gamma}$

egamma fake rate Data and MC

► Data

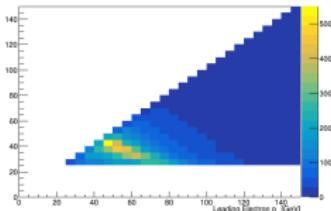


► Monte Carlo

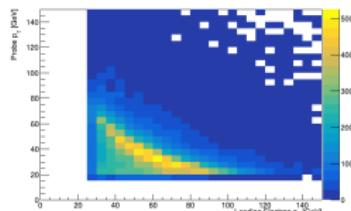


$e \rightarrow \gamma$ Data and MC Distributions

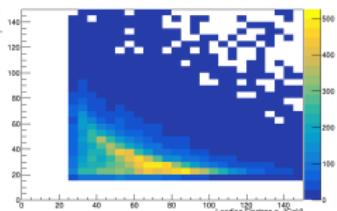
Data



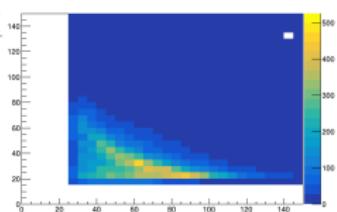
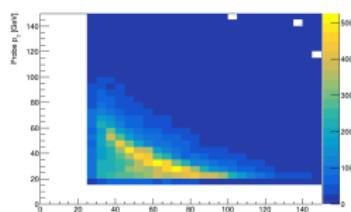
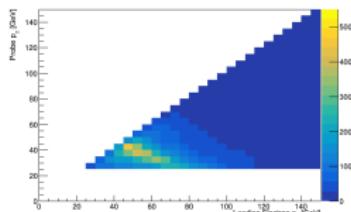
► Converted γ



► Unconverted γ



MC



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
- ▶ Optimizer: Adam
- ▶ Loss Function: Binary Cross Entropy

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

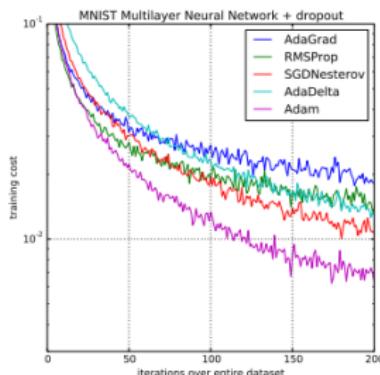
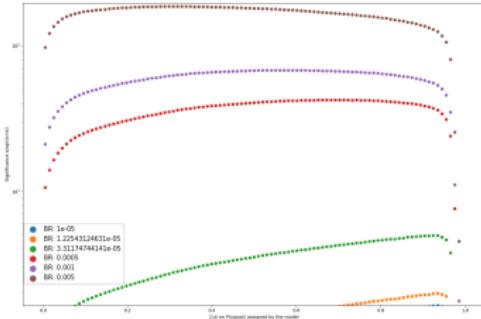
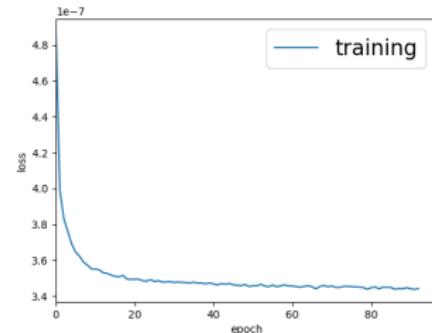
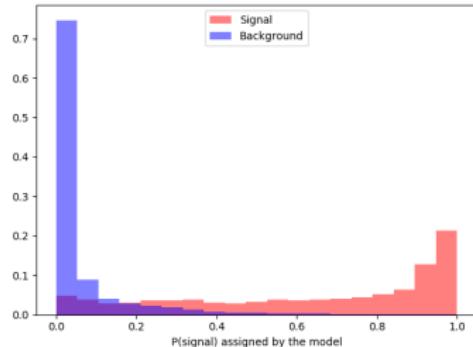
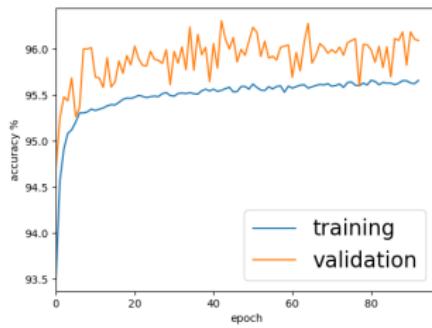


Figure: [Ref:Machine Learning Mastry]

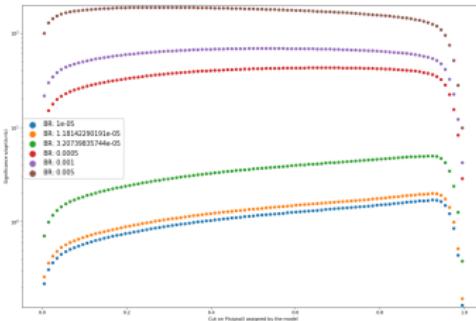
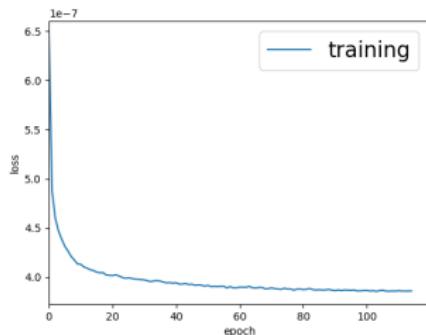
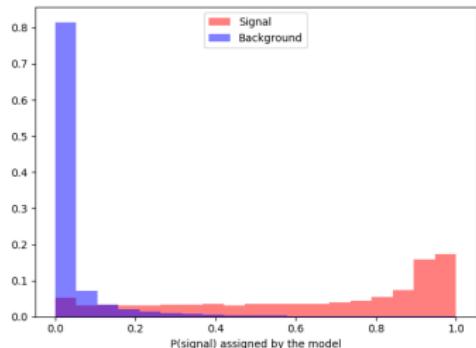
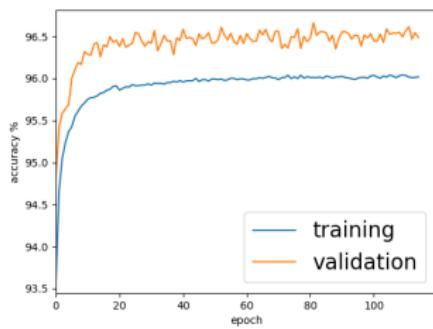
Neural Network Outputs

e+jets Channel Example



Neural Network Outputs

mu+jets Channel Example



MWT:

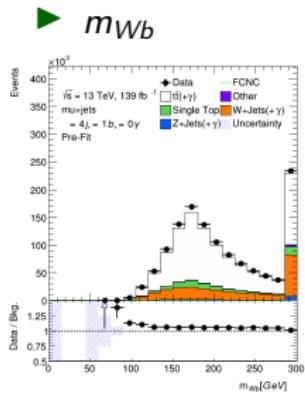
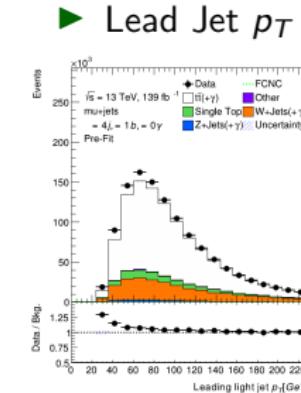
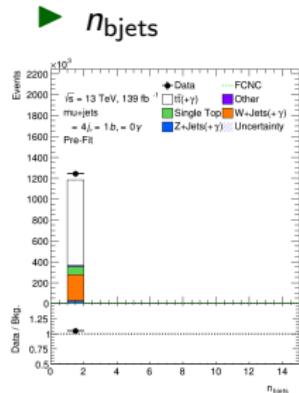
$$m_T^W = \sqrt{2p_{Tl} \times MET \times (1 - \cos(\Delta\phi))}$$

tqgam 6th order lagrangian:

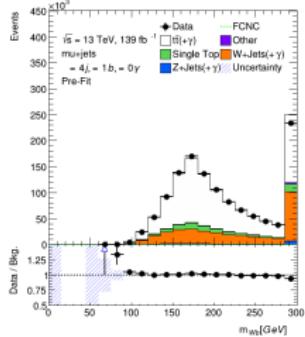
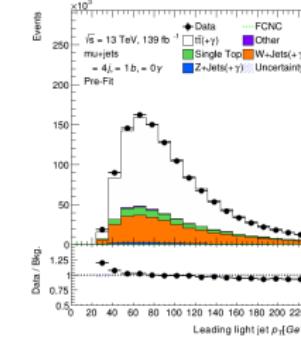
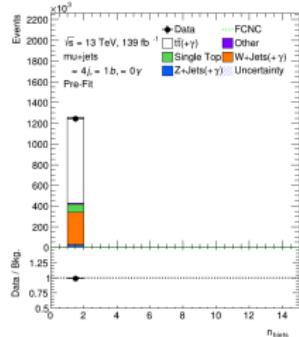
$$\mathcal{L}_{tq\gamma}^{eff} = -e\bar{c}\frac{i\sigma^{\mu\nu}q_\nu}{m_t}(\lambda_{ct}^L P_L + \lambda_{ct}^R P_R)tA_\mu + H.c.$$

0 Photon Regions: VR before/after SF, muon channel

Before SF



After SF



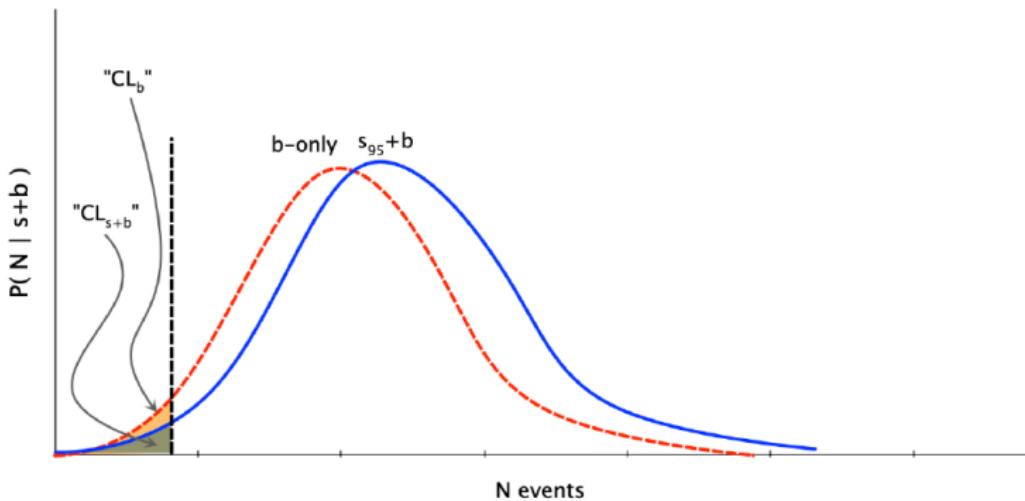
To address the sensitivity problem, CLs was introduced

<http://inspirehep.net/record/599622>

- common (misused) nomenclature: $CL_s = CL_{s+b}/CL_b$
- idea: only exclude if $CL_s < 5\%$ (if CL_b is small, CL_s gets bigger)

CL_s is known to be “conservative” (over-cover): expected limit covers with 97.5%

- Note: CL_s is NOT a probability



Cranmer