INSTITUTE FOR COLLIDER PARTICLE PHYSICS
UNIVERSITY OF THE WITWATERSRAND

ATLAS EXPERIMENT

Top-quark background estimation for physics BSM in the di-lepton and jets final state with the ATLAS detector

Kgomotso Monnakgotla

jeremiah.kgomotso.monnakgotla.cern.ch

School of Physics and Institute for Collider Particle Physics
University of the Witwatersrand

SAIP Conference 2019 Polokwane

July 11, 2019

Content

(Some) ATLAS and CMS Results

Motivation
Multi-lepton anomalies at the LHC

reference: **arXiv:1901.05300**

A wide range of analysis showing excesses in the multilepton production at the LHC has been studied.

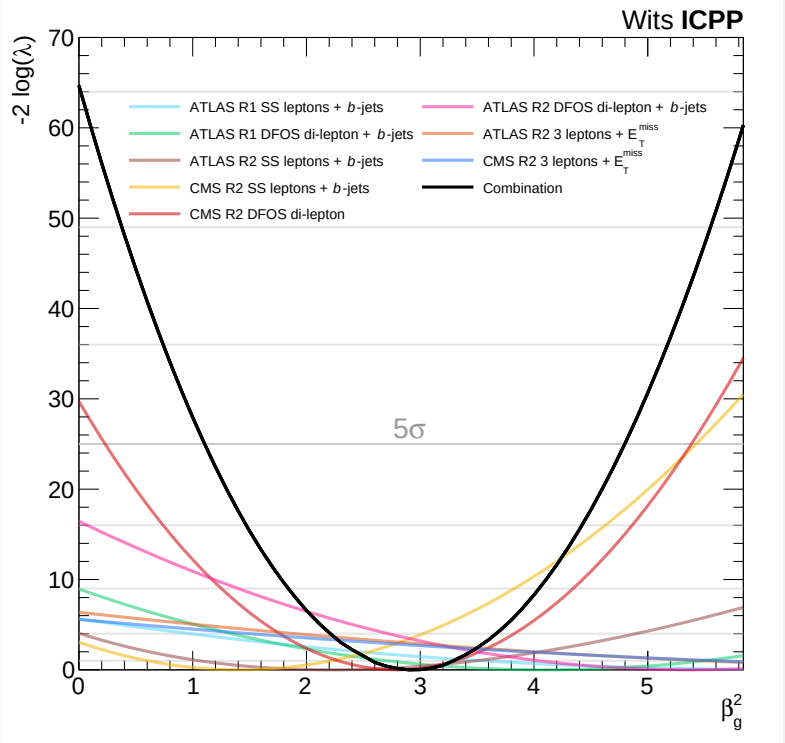
The profile likelihood ratios for each of the individual fit results with that of their combination

Results

– The combined excesses corresponds to a significance of 8σ

– These excesses can be explain by the Madala Hypothesis

– it is clear that di-lepton data is not well understood

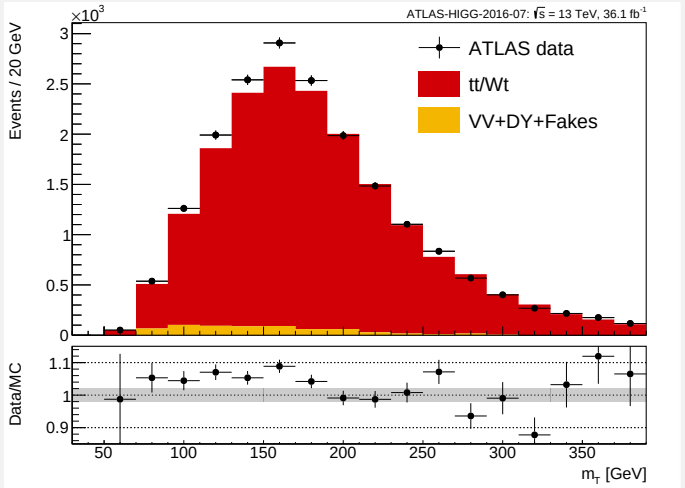


Motivation
Top quark measurement in CMS

Measurement of lepton differential distributions and the top quark mass with CMS @13 TeV: **JHEP 10 (2018) 117**

Analysis Selection

- Select dileptonic $t\bar{t}$ events: $e^{\pm}\mu^{\mp}$ final state
- Inclusive analysis: no requirement on jets!

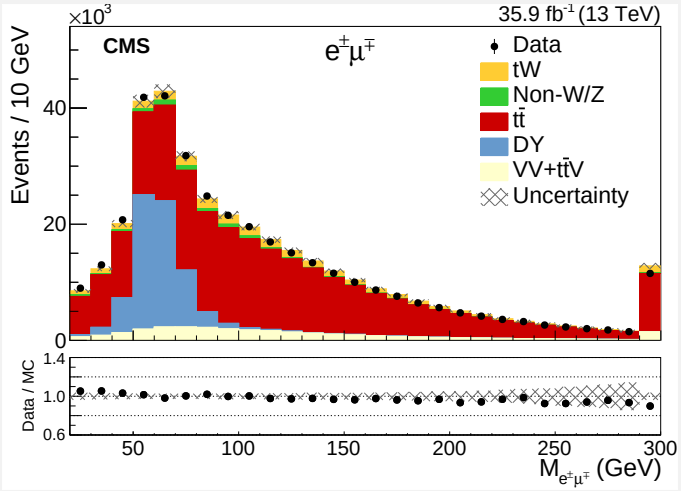


Results

- m_T distribution well described at high mass region
- High discrepancy for m_T at the peak
- discrepancy at low $m_{\ell\ell}$

* Is this a top MC mismodeling?

* **Are there other non-top results observing these features?**



Motivation

Measurement of the WW XSEC in ATLAS

ATLAS @ 8 TeV: **JHEP 09 (2016) 029**

ATLAS @13 TeV: **arXiv:1905.04242**

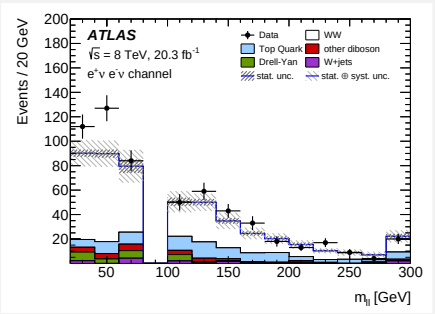
Analysis Selection

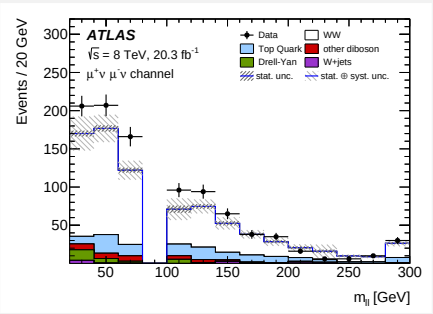
- Select $\ell^{\pm}\ell^{\mp}$ and **vetoing** events with jets
- ee / $\mu\mu$ (top-row)
- $e\mu$ (middle and bottom row)

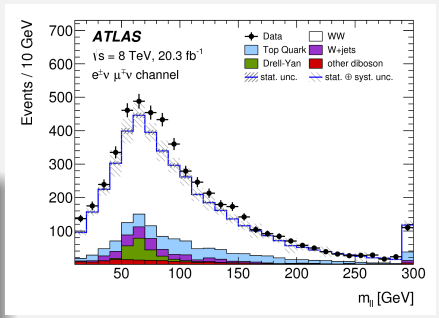
Results

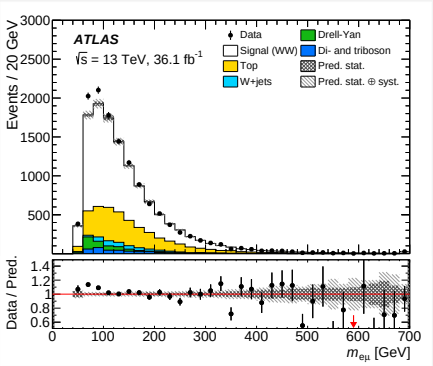
- High discrepancy for $m_{\ell\ell} < 100$ GeV
- High discrepancy for lepton p_T around 100 GeV
  - very unlike to be produced by fake leptons

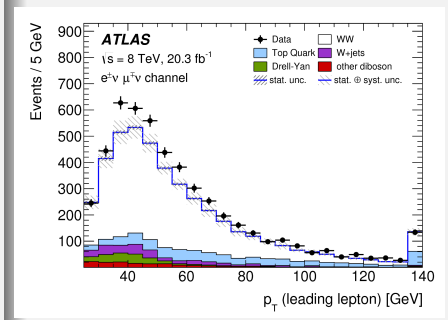
* **New physics or poor modelling?**

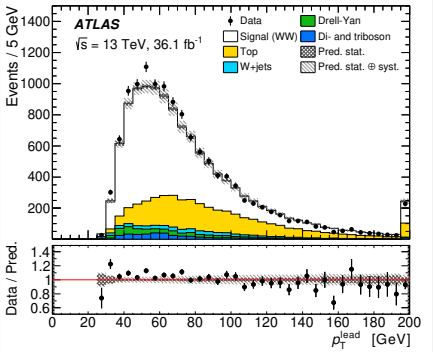












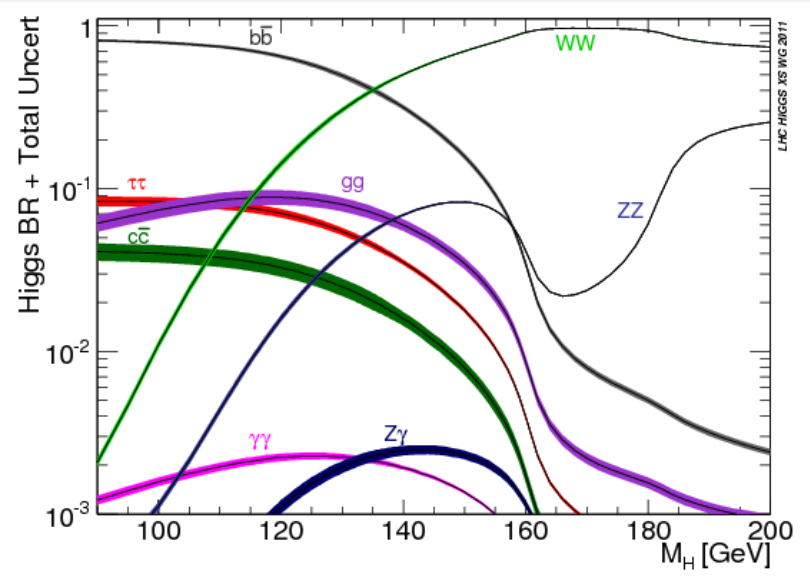
The Madala Hypothesis

The Madala Hypothesis

References: **Eur. Phys. J. C (2016) 76: 580**, **J.Phys. G45 (2018) no.11, 115003**, **arXiv:1809.06344**, **arXiv:1901.05300**

$gg \rightarrow H \rightarrow Sh$

- H : Heavy Scalar with mass range [240,2m$_t$] GeV
- S : Higgs-like Scalar with between m_h and m_H
- h : Standard Model (SM) Higgs boson with mass m$_h$ = 125 GeV



S boson: BRs decays depend on its mass

- $>$ **135 GeV:** $S \rightarrow WW$ becomes the dominant decay

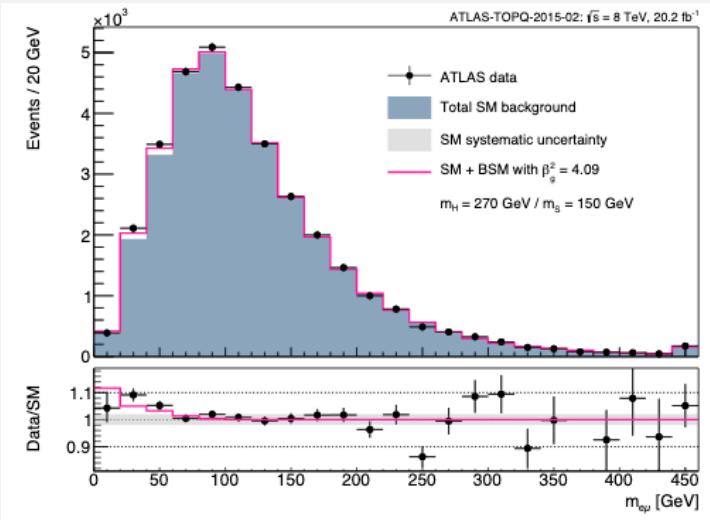
$H \rightarrow Sh$ produces a wide range of possible final states, dominated by leptons and jets

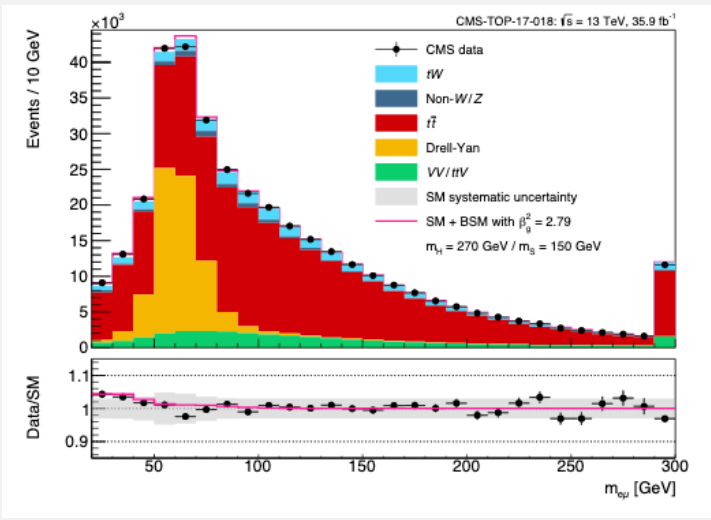
Fits to LHC data

The SM+BSM fit results for the di-lepton invariant mass spectrum from ATLAS in Run 1 (left) and CMS in Run 2 (right).

Results

- It is clear that there exists an excess of events at $m_{\ell\ell} < 100$ GeV
- Fit the $H \to Sh$ model to CMS data and ATLAS data in the region where $m_{\ell\ell} < 110$ GeV
- Fit of data results from **arXiv:1901.05300**





Object Selection

The **final state** for this search is characterized by **two** oppositely charged **leptons** and **jets**

Requirement for Leptons

Electrons:

- $p_T^e > 15$ GeV
- $|\eta| < 2.47$ excluding $1.37 < |\eta| < 1.52$

Muons:

- $p_T^\mu > 15$ GeV
- $|\eta| < 2.5$

Requirement for Jets

Jets:

- $p_T > 25$ GeV
- $|\eta| < 2.5$

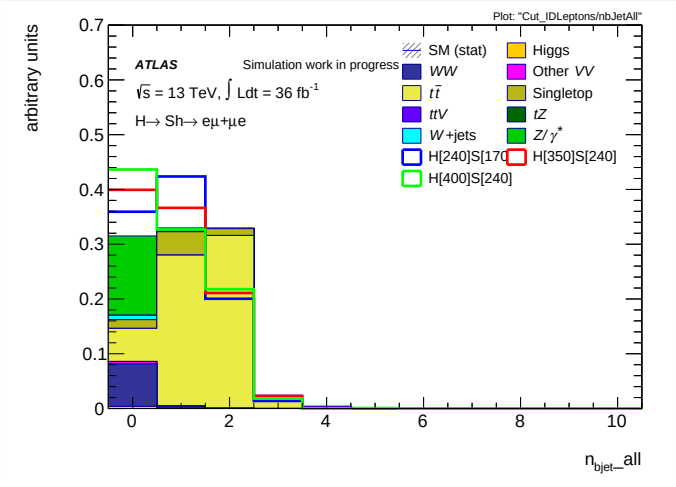
b-tagged jets:

- MV2C10 @ 85 %
- $p_T > 20$ GeV

Select events with leading lepton $p_T > 27$ GeV, Sub-leading lepton $p_T > 15$ GeV and $\geq$2 jets

Signal Characterization: Jet multiplicities

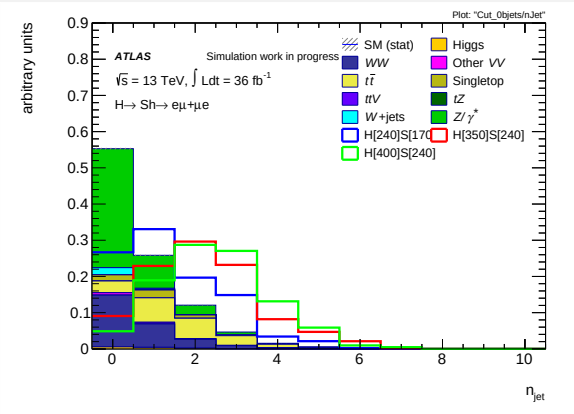
Number of b-tagged Jets ↓



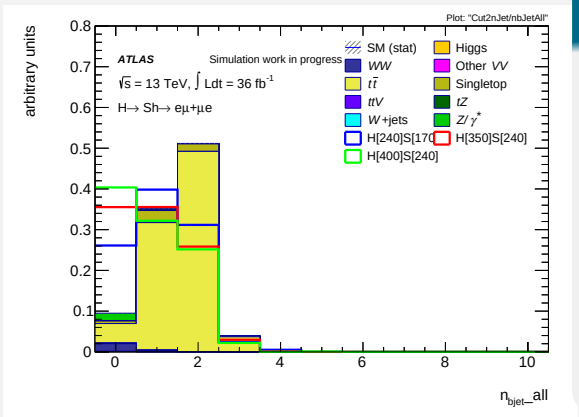
Signal Features

- $N_{(b)jet}$: Veto events with *b*-jets: **Reduce** Top background
- Select events with $\geq$ 2 jets: High Signal/Background

Number of Jets after *b*-jets veto ↓



↓ After $\geq$ 2 N_{jets}



Highest Signal/Background

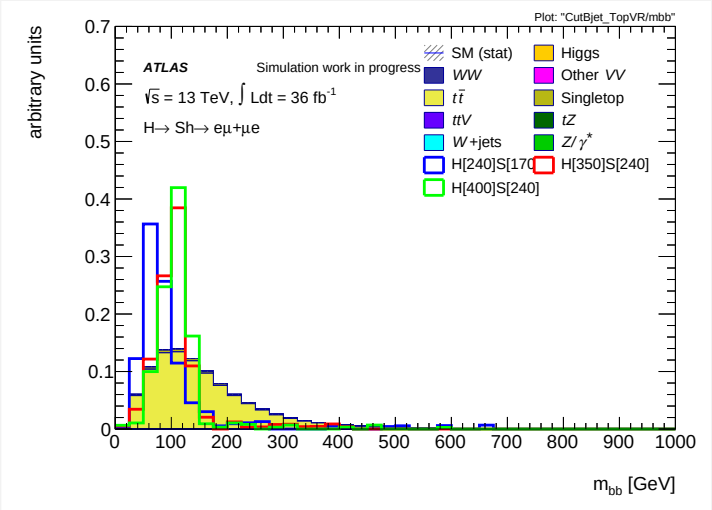
- $N_{jets} \geq$ **2+ *b*-jet veto**
- Dominanted by Top backgound
- top purity: 60 %

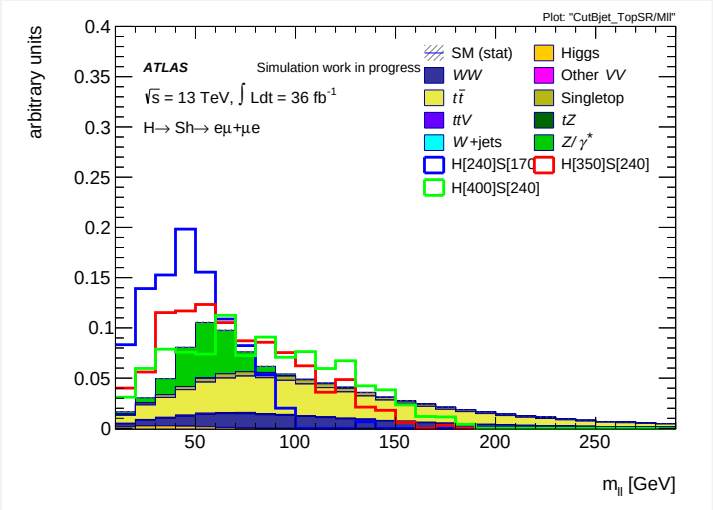
Top background estimation

Event Selection

The Top CR and VR are both defined to be orthogonal to the signal region selection and required to be signal depleted

Observable	SR	Top-VR	Top-CR
Leading lep p_T	$>$ 27 GeV		
Sub-leading lep p_T	$>$ 15 GeV		
Number of jets	$\geqslant$2		
Number of b-tagged jets	=0	=2	=0
m_{bb}	-	>150 GeV	-
$m_{\ell\ell}$	-	-	$>$ 150 GeV



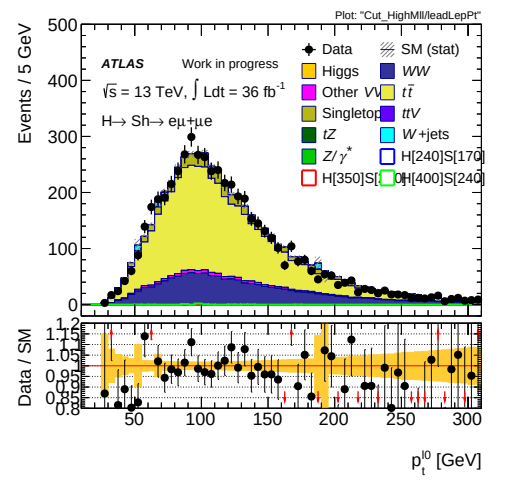


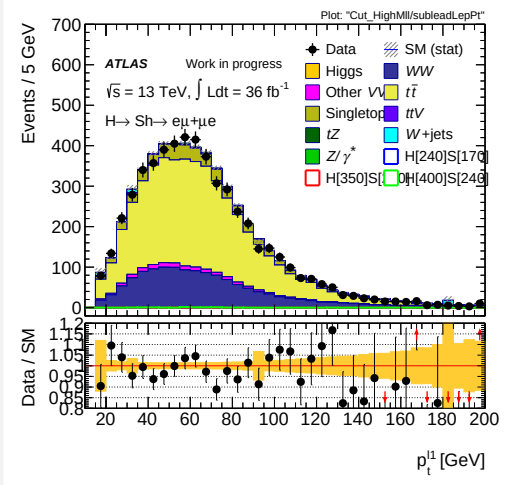
Top Control Region: 0-*b*jets+$m_{\ell\ell}$>150 GeV

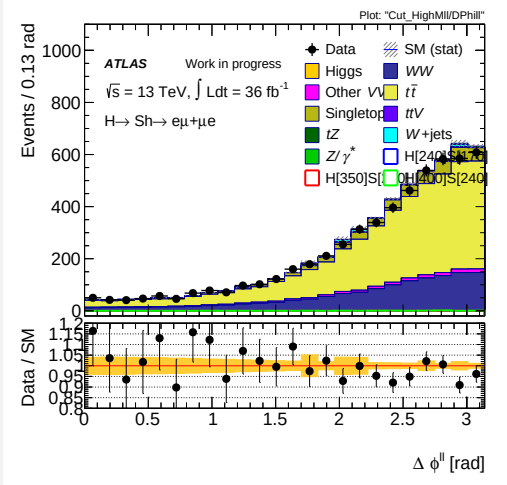
	Higgs	WW	Other VV	$t\bar{t}$
0 b-jets+$M_{\mathrm{ll}} > 150$ GeV	3.19 ± 0.08	1355.89 ± 7.01	146.37 ± 1.90	3534.73 ± 17.30

Singletop	ttV	W+jets	Z/γ^*	Total Bkg	Data/bkg	data
469.49 ± 13.34	9.46 ± 0.26	40.44 ± 20.30	18.01 ± 10.83	5604.98 ± 31.56	0.97 ± 0.01	5456

- Define Top CR by 0-*b*-jet + $m_{\ell\ell} >$ 150 GeV
- top-quark is the dominant background, mostly $\mathbf{t\bar{t}}$
- data/MC agreement : 0.97
- Top($t\bar{t}$ + Wt) purity: **72**%



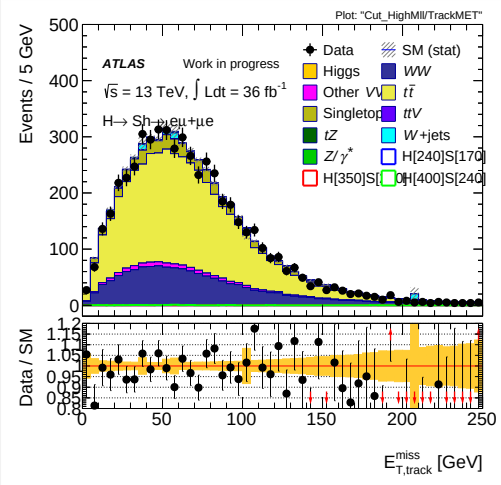


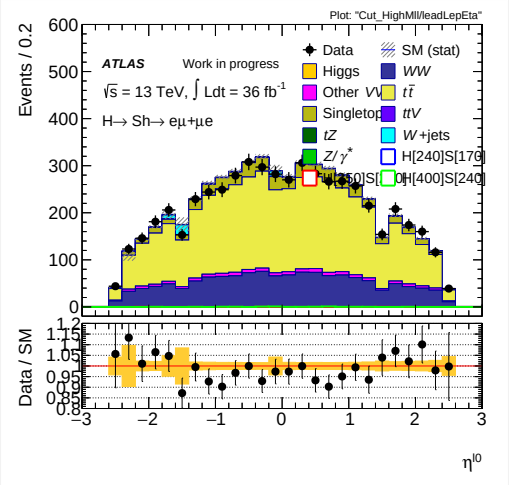


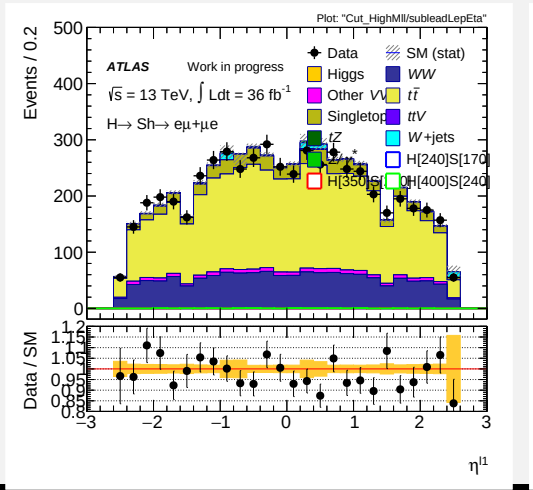
* p_t^{l0}: Leading lepton p_T ↑ p_t^{l1}: Sub-leading lepton p_T ↑

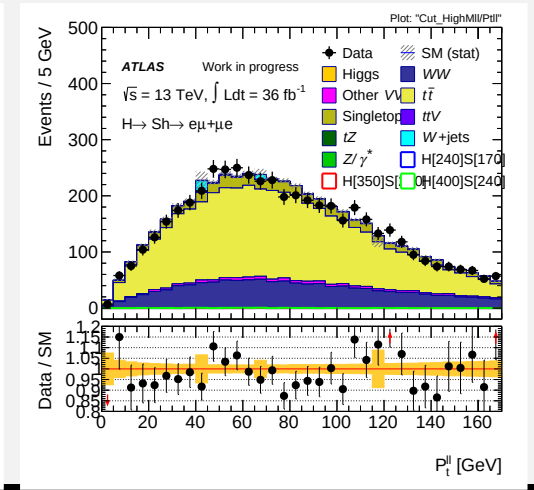
Top Control Region: 0-*b*jets+$m_{\ell\ell}$>150 GeV

- The contribution from non-resonant *WW* (**see Lebohang Mokoena's talk**) is about 25%
- The contribution from *W*+jets, other *VV* and Z/γ^* processes add up to 3%
- Good agreement data/MC: **0.97**
- The residual mismodelling is due to missing EW corrections









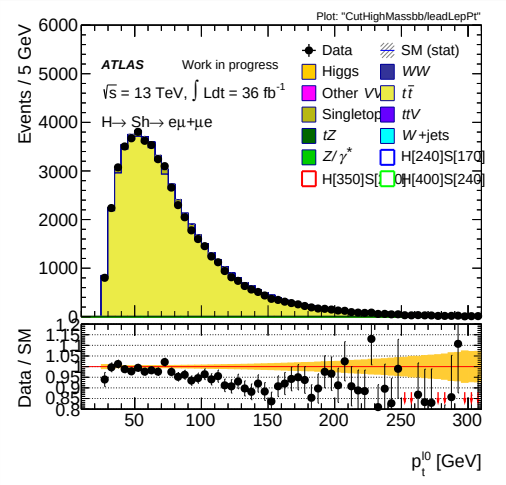
Top Validation Region: 2-*b*jets+m_{bb}>150 GeV

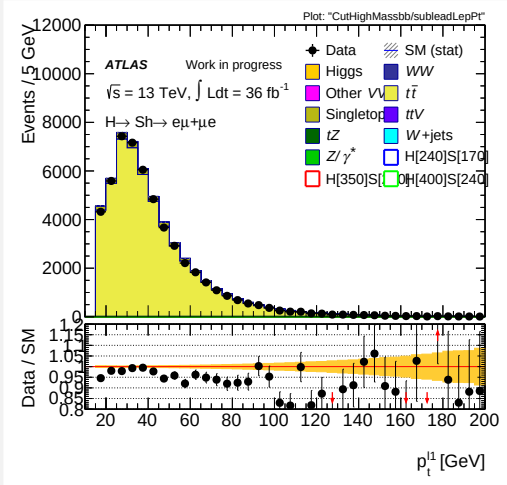
	Higgs	*WW*	Other *VV*	$t\bar{t}$
2 *b*-jets+$M_{\mathrm{bb}} > 150$ GeV	25.29 ± 0.24	33.05 ± 0.84	6.43 ± 0.33	52658.56 ± 64.70

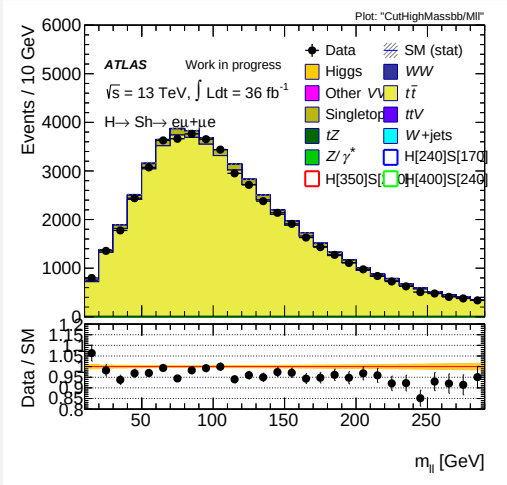
Singletop	*ttV*	*W*+jets	Z/γ^*	Total Bkg	Data/bkg	data
2134.64 ± 17.98	85.25 ± 0.75	0.00	62.93 ± 3.76	55011.99 ± 67.29	0.97 ± 0.00	53108

- Define Top VR by 2-*b*-jet + $m_{bb} > 150$ GeV
- top-quark enriched region

- data/MC agreement : 0.97
- Top($t\bar{t}$ + *Wt*) purity: **99**%
- The MC residual mismodelling is due to missing EW corrections

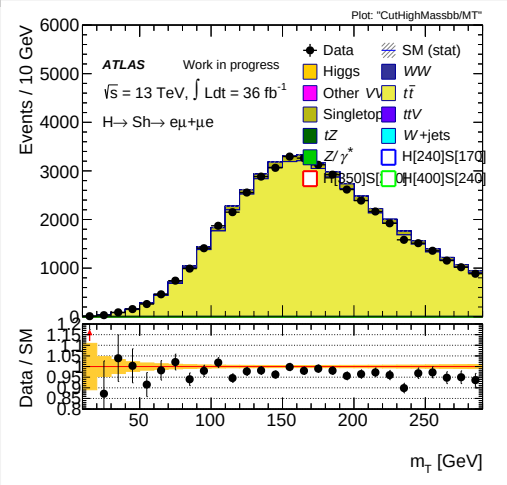


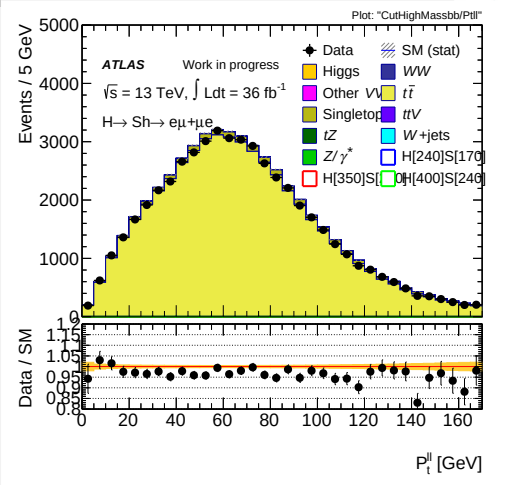


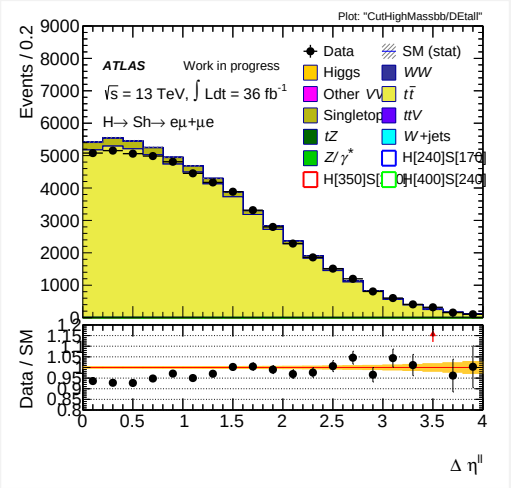


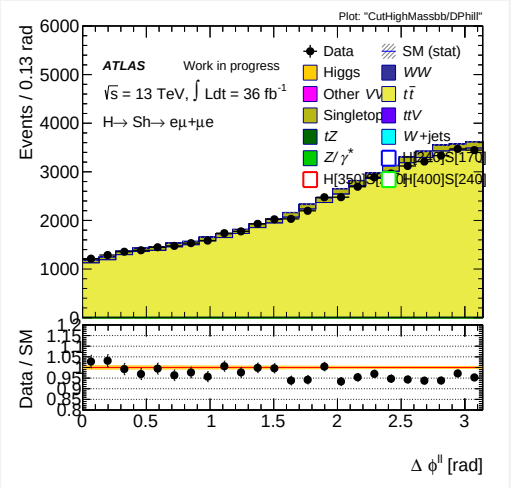
Top Validation Region:2-*b*jets+m_{bb}>150 GeV

- Other Background contributions add up 1%
- Good agreement data/MC: **0.97**
- m_T and $p_T^{\ell\ell}$ are well described by the MC
- Observe a MC overshoot at low $\Delta\eta^{\ell\ell}$and high $\Delta\phi^{\ell\ell}$









Summary

- Several results from ATLAS and CMS observes discrepancy between MC and data
- The Madala hypothesis allows to better describe the data in top and WW SM results at LHC
- The search for Madala Hypothesis is performed with two opposite charged leptons and jets in the final state
- top background is the dominant SM process to the $H \rightarrow Sh$ search so it is crucial to validate the top MC
- Top-CR has good data/MC agreement of 0.97 with 72% top purity
- The top-VR has good data/MC of 0.97 with 99% top purity
- The MC works very well in the regions of the phase-space where the signal is not expected

Thank You!