



New dark matter searches with E_{miss}^T + bb final states at ATLAS

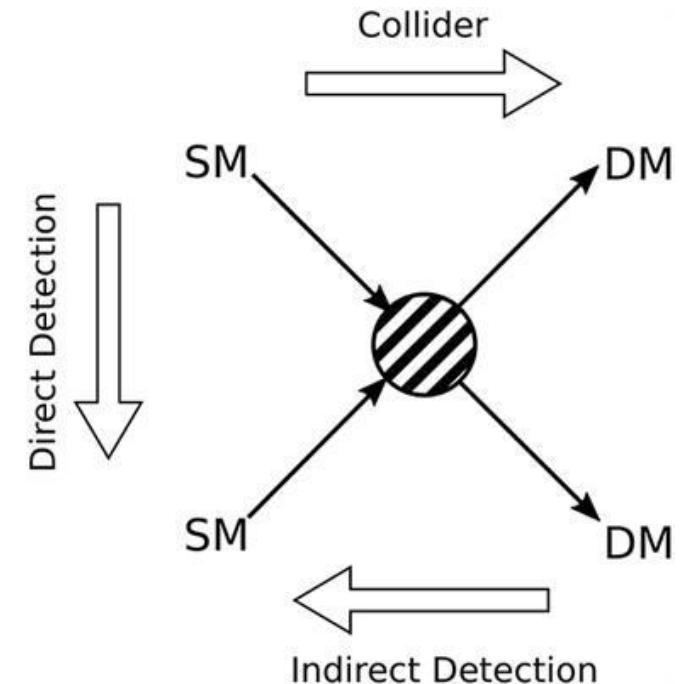
Candice Basson
IoP Joint APP, HEPP and NP conference
April 2021



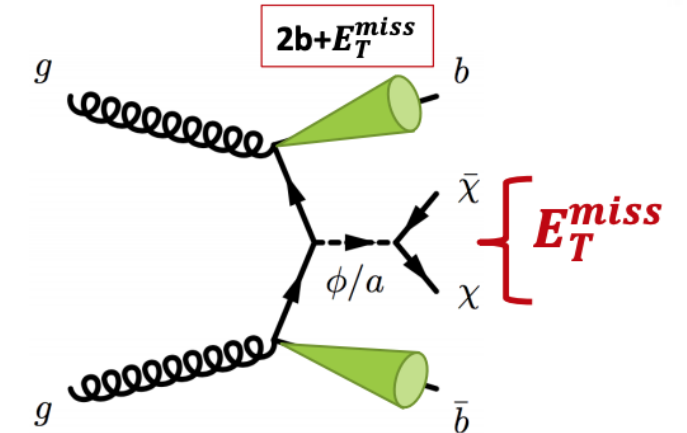
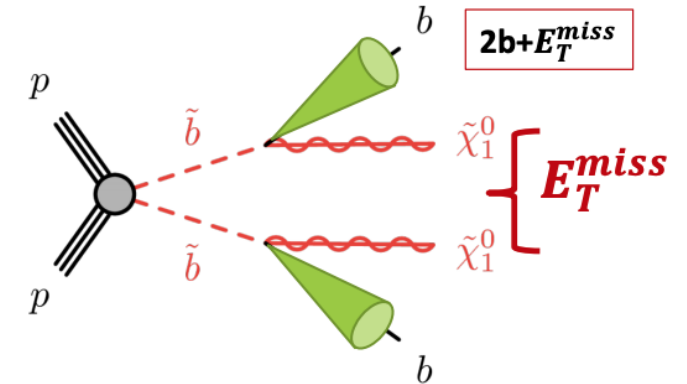
MANCHESTER
1824

The University of Manchester

- Dark matter (DM) may be composed of weakly interacting massive particles (WIMPs)
- WIMPs could be produced in pairs at the LHC via the decay of a new mediator particle that couples to SM quarks
- The lightest supersymmetric particle (LSP) is a DM WIMP candidate



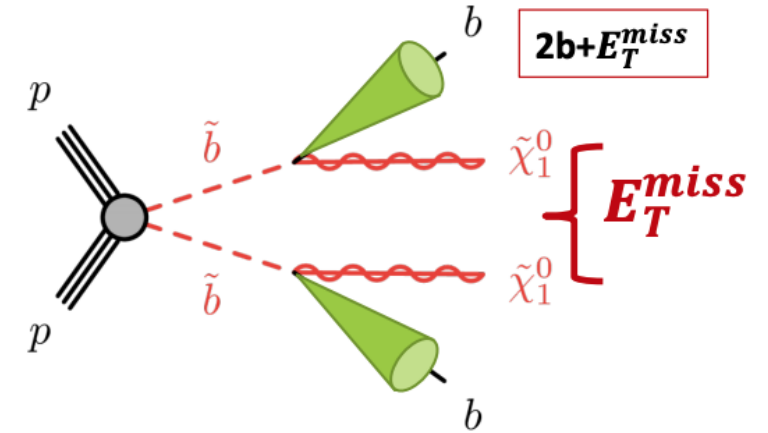
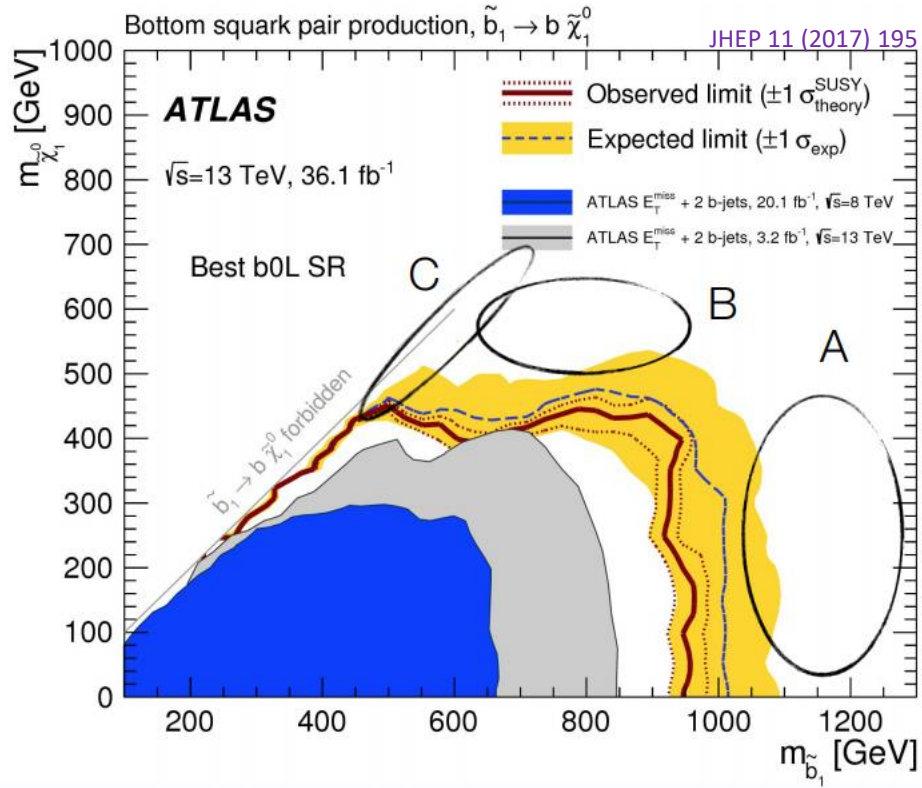
- Analysis searching for new physics in final states with 0 lepton, 2 b-jets and missing transverse energy (MET)
- Analysis using full run-2 139fb^{-1} dataset
 - this iteration builds on two of the 36.1fb^{-1} [sbottom](#)¹ and [DM+bb](#)² analyses due to a shared final state
 - Improvements in analysis strategy, such as the employment of machine-learning techniques and soft b-tagging
- Public as of 1st Feb 2021 [\[arXiv:2101.12527\]](#)



[1] arXiv:1708.09266

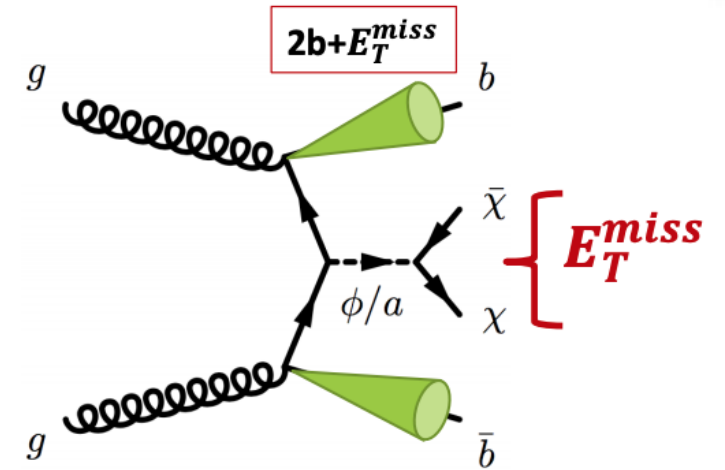
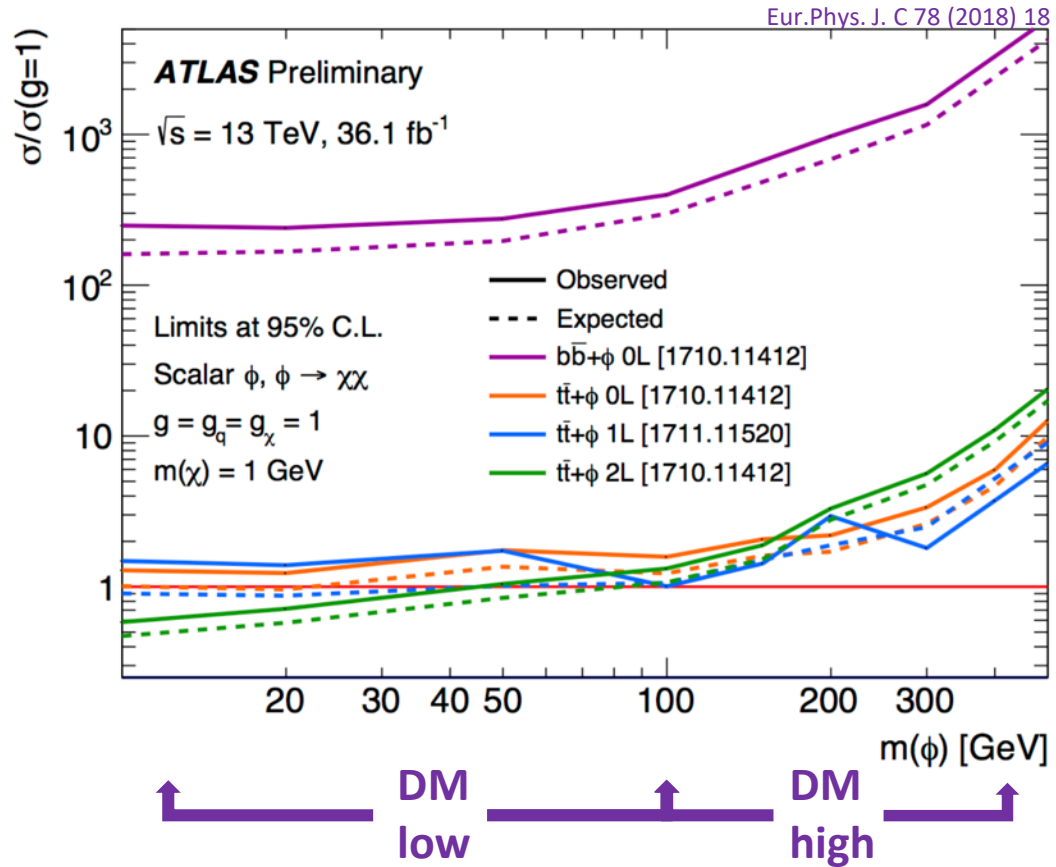
[2] arXiv:1710.11412

- Sbottom pair production, each decaying to a b-jet and a neutralino (LSP) with BR = 100%

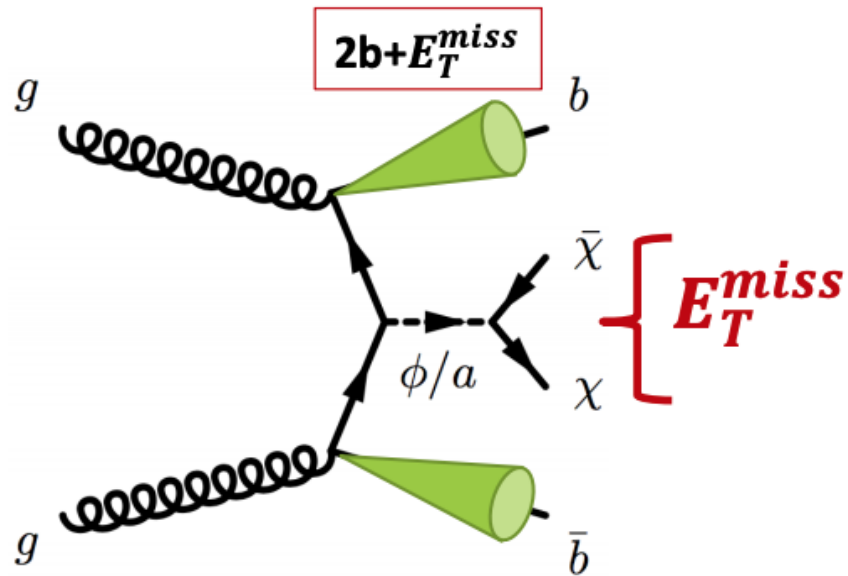


- The sbottom/ neutralino mass plane mass plane is targeted with 3 fiducial regions
 - bulk region – **A**
 - Intermediate region – **B**
 - Compressed region - **C**
- Now combined into AB to increase coverage

- DM production in association with b-jets, with either a scalar (ϕ) or pseudo-scalar (a) mediator to the dark sector



- $b\bar{b} + \phi$ 0L behind $t\bar{t} + \phi$ exclusions due to less sensitivity
 - lots of room to work with!
- Split into two regions
 - low mass mediators – **DM_low**
 - high mass mediators – **DM_high**



Optimized to target DM by making use of a BDT
 $\cos(\vartheta)^*$ main discriminating variable

$$\cos(b_1, b_2)^* = \left| \tanh\left(\frac{\Delta\eta_{bb}}{2}\right) \right|$$

Low mediator mass:
 10-100 GeV

- invariant bb mass peak
- lower MET
- higher angular separation of b-jets

High mediator mass:
 100-500 GeV

- peak broadens in invariant b mass
- higher MET
- less angular separation of b-jets

Why is $\cos(\vartheta)^*$ a useful variable for DM discovery?

- this variable is expected to be mostly flat for SM but in general shows enhancement in high $\cos(\vartheta)^*$ for high mass scalar mediators and in all pseudo-scalars mediators

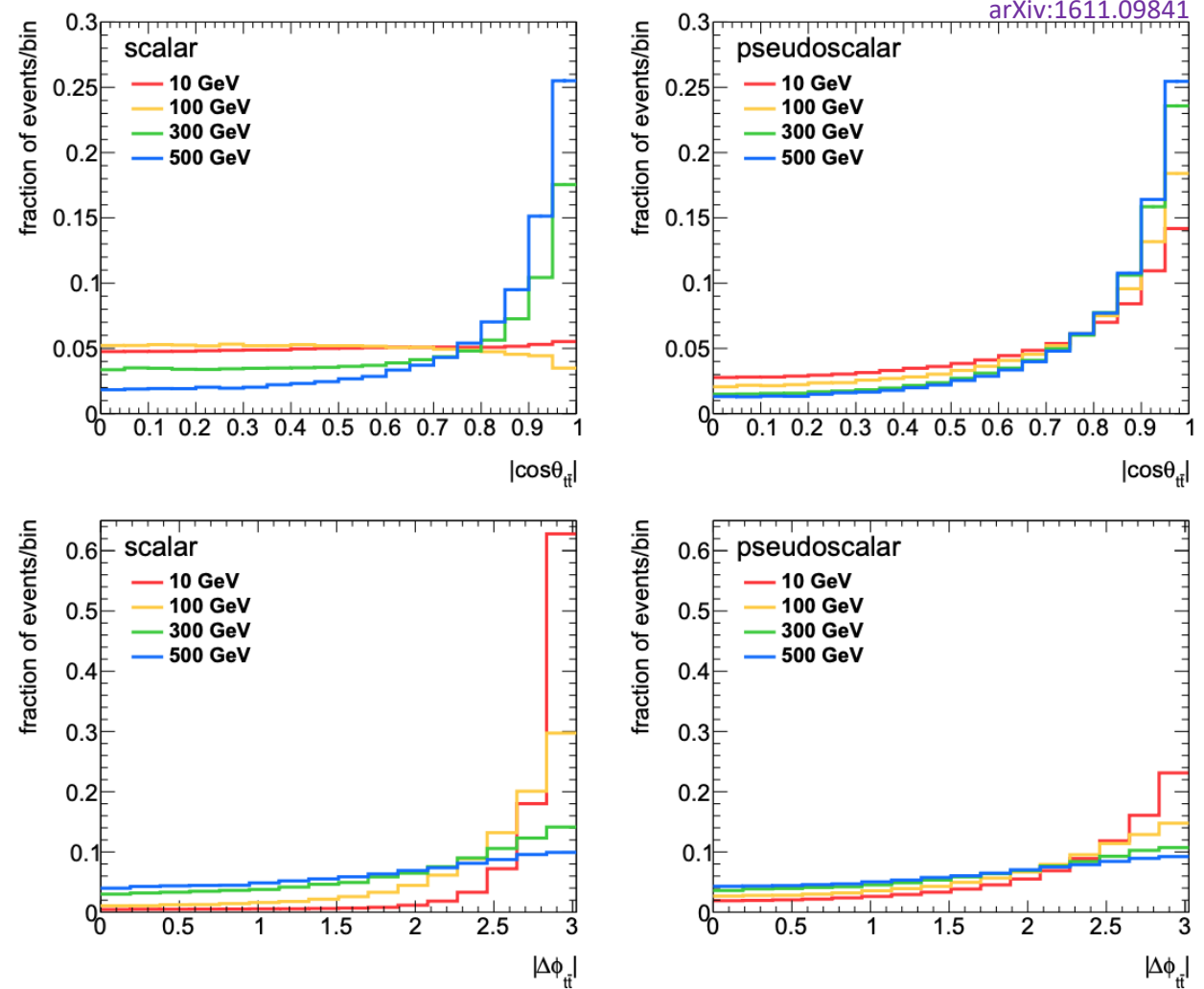
- This is discussed more in detail in a paper [\[arXiv:1611.09841\]](https://arxiv.org/abs/1611.09841)

- These plots also show why the delta variables are useful in the BDT

$$\delta^+ = |\Delta\Phi(E_T^{\text{miss}}, j^{1-3}) + \Delta\phi_{bb} - \pi|$$

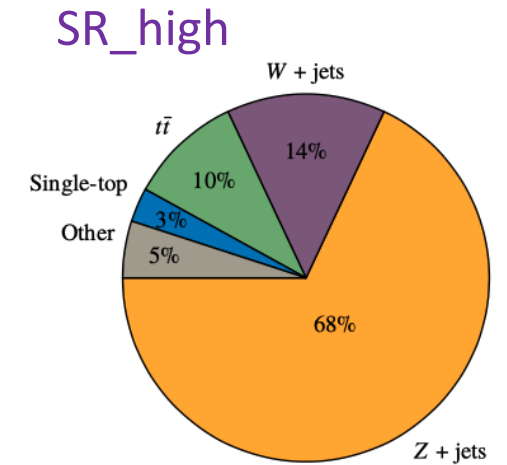
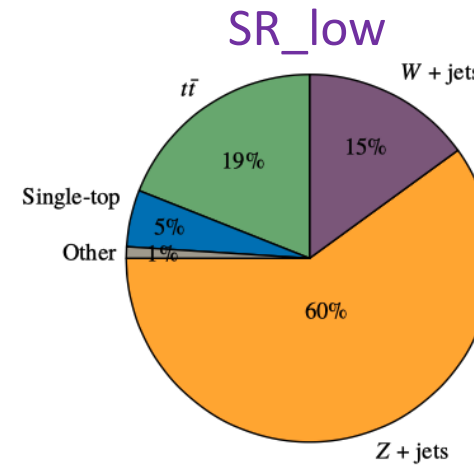
$$\delta^- = \Delta\Phi(E_T^{\text{miss}}, j^{1-3}) - \Delta\phi_{bb}$$

N.B $\Delta\phi_{bb}$ is small for the $Z(\nu\nu)+bb$ background from the gluon splitting process



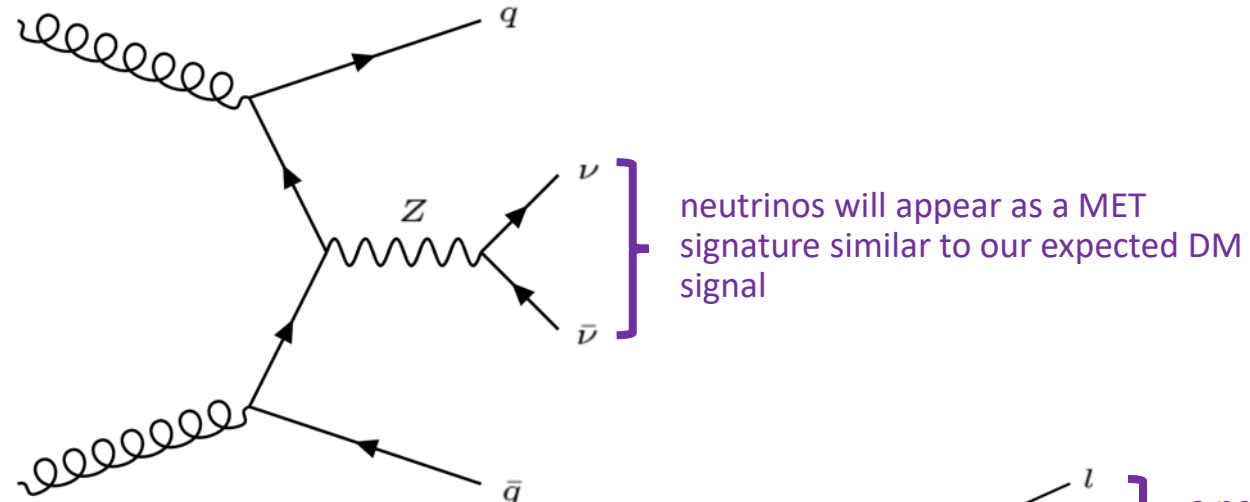
arXiv:1611.09841

- Two BDTs trained on either low or high mediator masses
 - input variables include leading jets' p_T , MET and angular variables related to the jets
- **Z+jets** are the dominant background in both of our signal regions followed by **W+jets** and **ttbar**

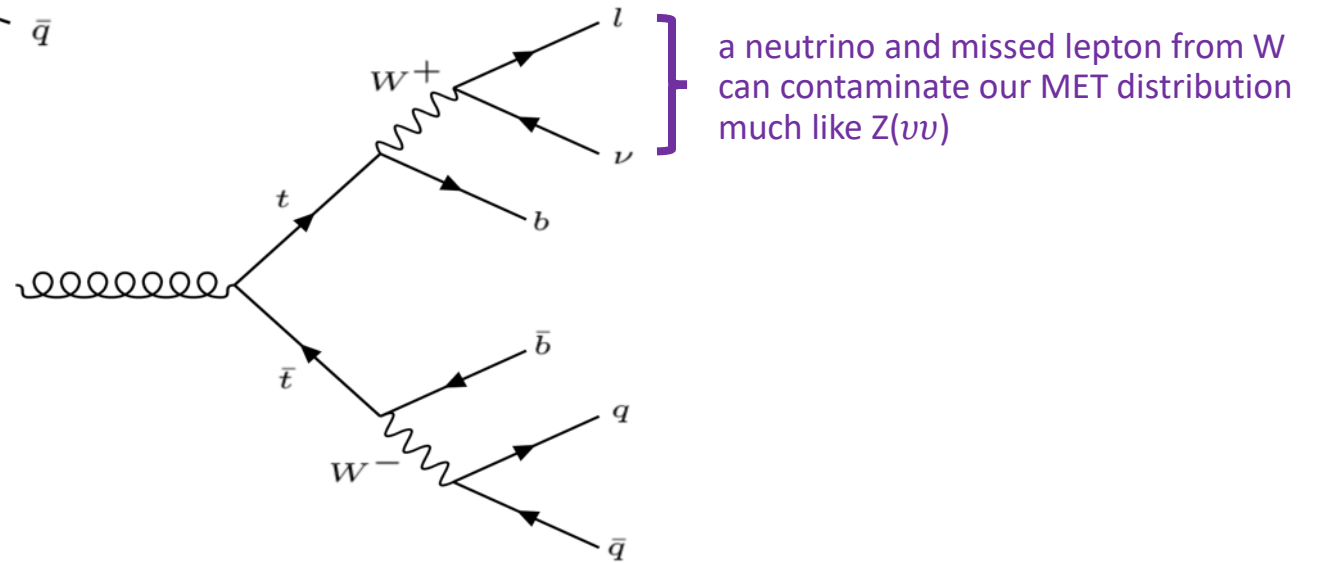


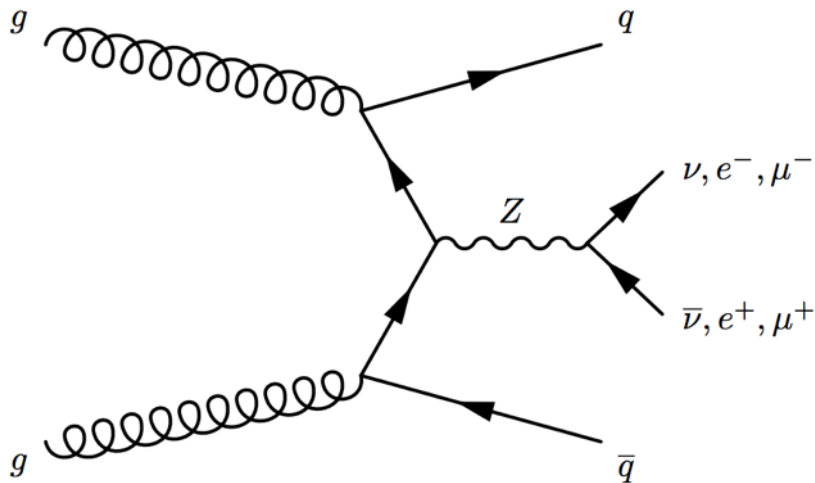
- Despite the large increase in data and the region being designed to enhance signal, the fraction of signal events is still small
 - e.g. last bin in $\cos(bb)^*$ for SR_low we expect about 145 SM events but only about 2 DM events (where the scalar mediator has a mass of 20 GeV)
- A signal free CR is needed to check modelling of SM backgrounds
 - Incorrect modelling can lead to an under or overestimate of the SM backgrounds which can enhance or mask signal events

Z backgrounds



top and W backgrounds





Z+jets CR:

- Use $Z \rightarrow ll$ which is kinematically similar to $Z \rightarrow \nu\nu$
- Mark leptons as invisible to replicate MET in signal
 - Introduces a Z p_T cut
- Cut on 'real' MET to remove $t\bar{t}$ and W backgrounds as $Z \rightarrow ll$ should not have a MET signature
- However this introduces additional selections
 - lepton selections
 - invariant mass of di-lepton system to select Z mass peak

Signal Region (SR):

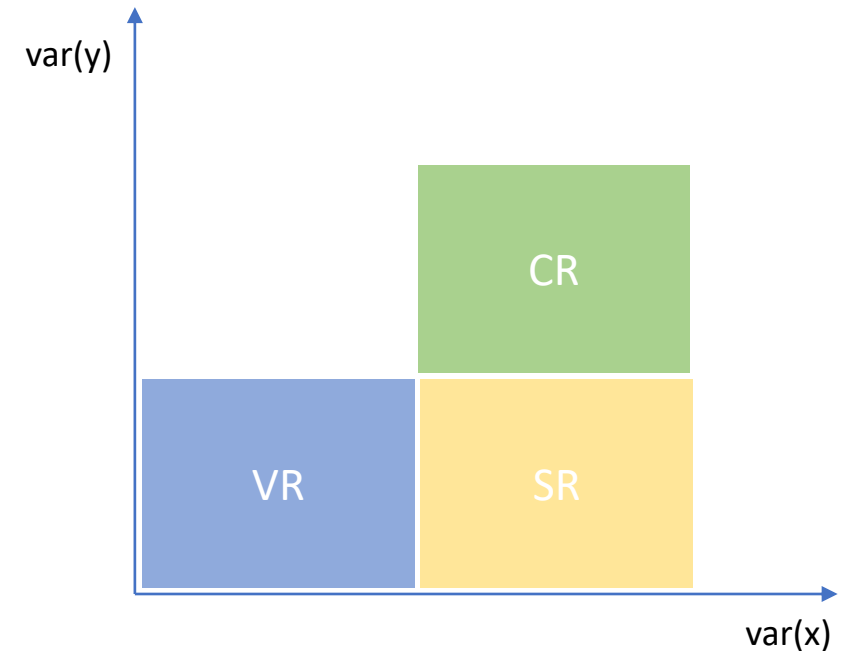
- A region designed to be high in signal and low in backgrounds

Control Region (CR):

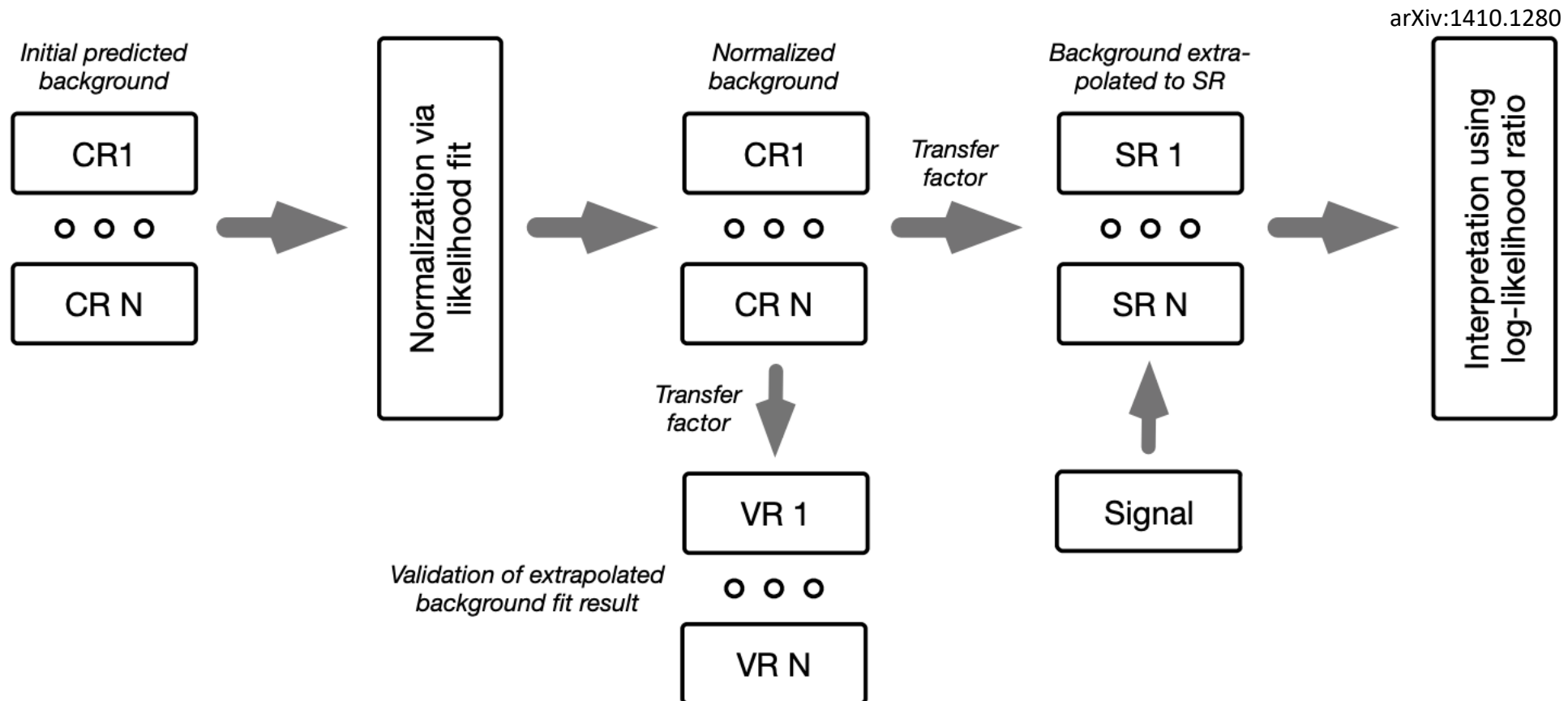
- A region designed to select a specific SM background
- Kinematically similar to the SR
- Used to check background modeling and derive a normalization factor through a background only fit
- No signal should be present

Validation Region (VR):

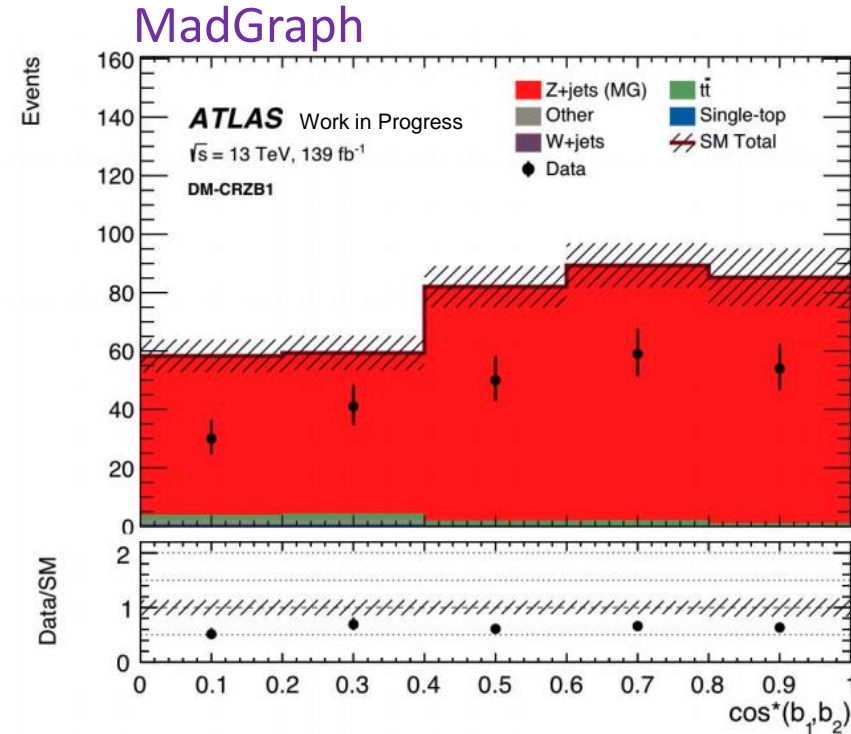
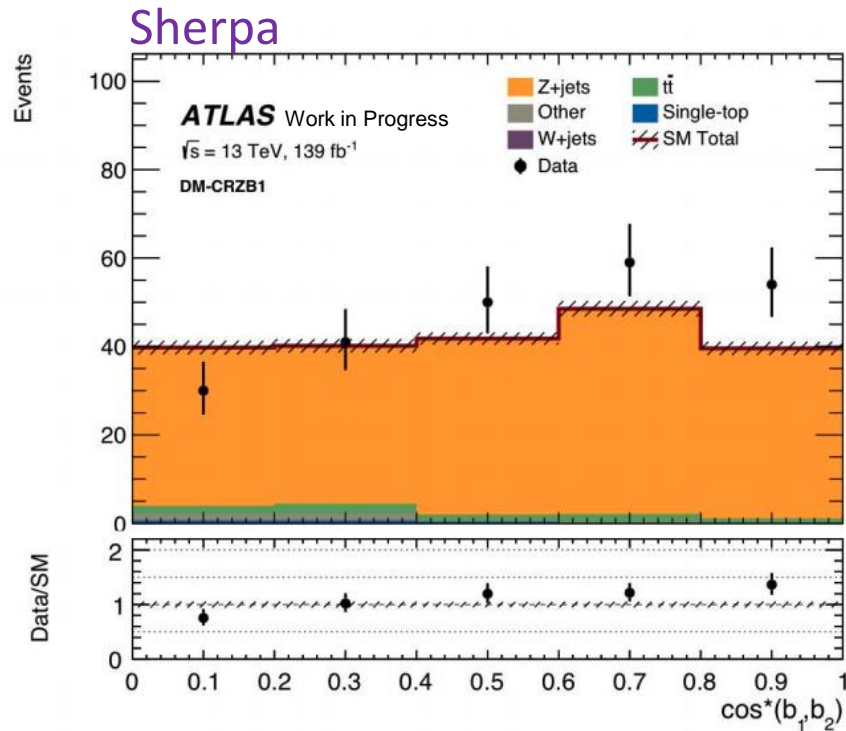
- Designed to validate the normalization factor derived from the CR
- Also kinematically similar to the SR



- A fitting framework (HistFitter) is used to perform a profile likelihood fit that uses:
 - the expected number of Monte Carlo (MC) events
 - the number of data events
 - the statistical and systematic uncertainties on the expected number of MC events
- A background only fit makes use of the Z CR to extract a normalization factor by fitting the Z Monte Carlo to data
- Other backgrounds are normalized using theoretical calculations

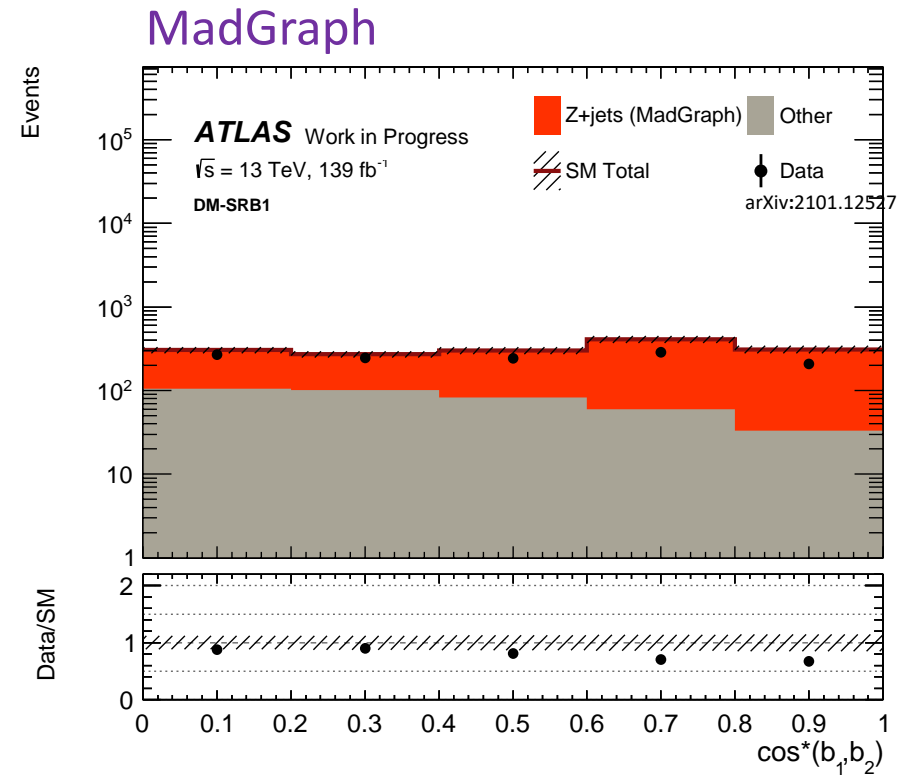
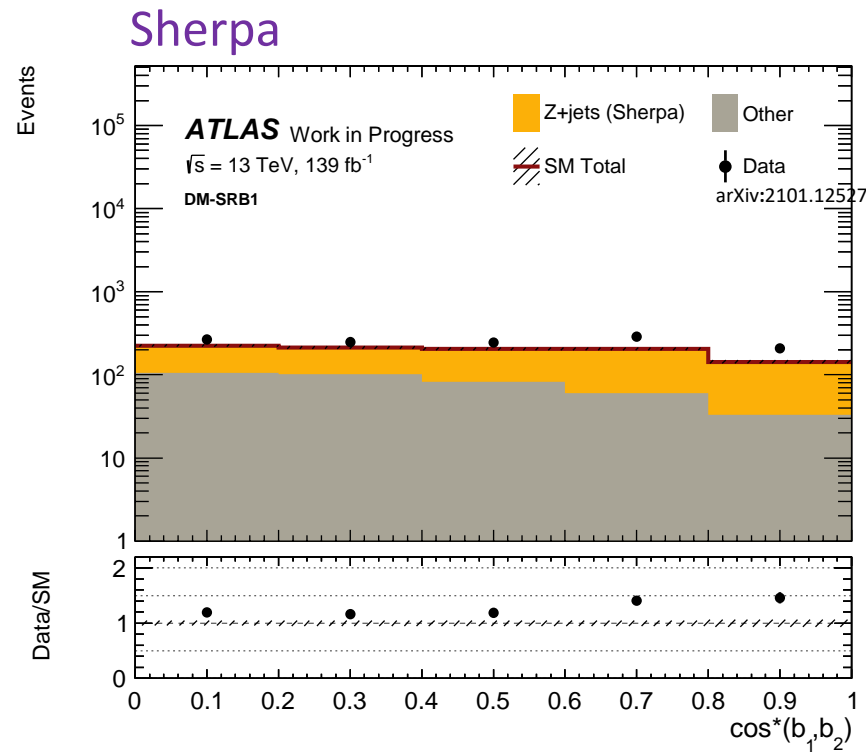


- A good CR is crucial, as significant differences in shape and normalization predicted between generators as shown below (pre-fit)
 - The same data and selections are used in both plots
 - Indications of problem in modelling and normalization that isn't 'solved' by either generator
 - Theory uncertainties within an MC prediction are not the only (or even main) source of uncertainty



N.B: Overall experiment and theory systematic uncertainties of $\sim 7\%$ not shown in plots above

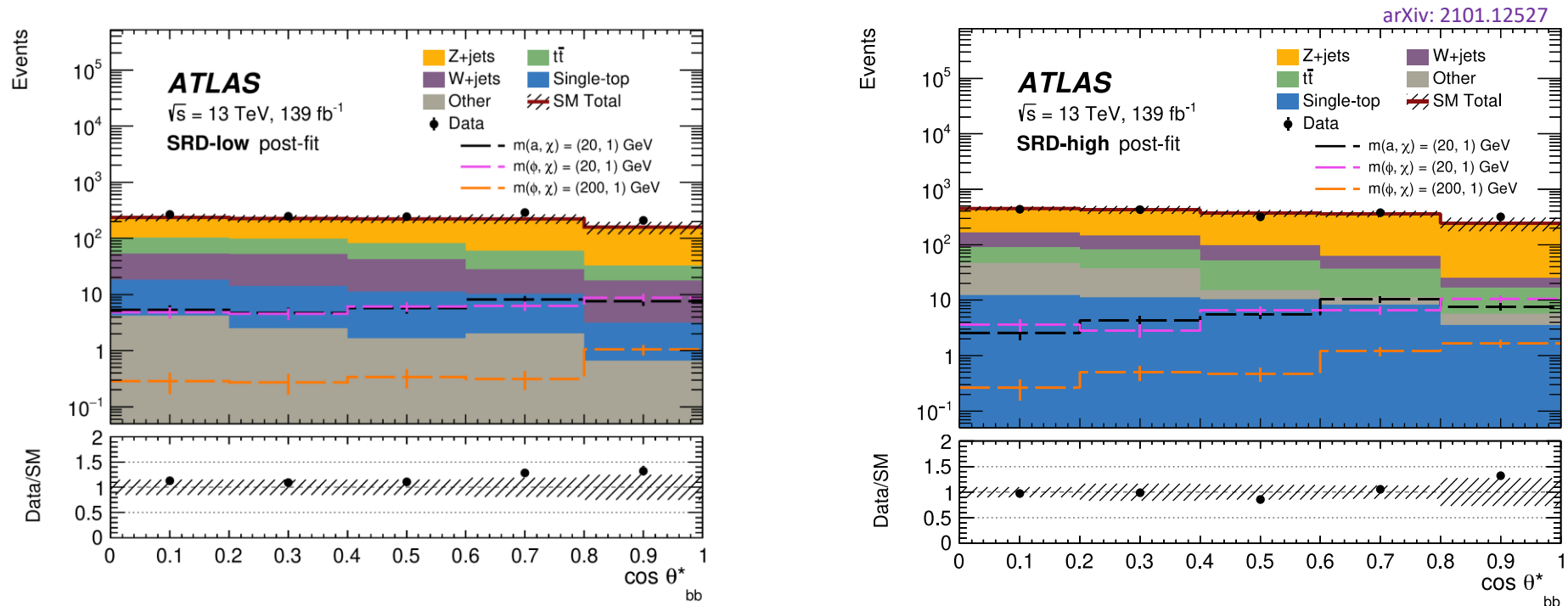
- Below plots show the $\cos(bb)^*$ distributions pre-fit for low mass mediators in the SR
- The difference in prediction from the different generators is seen to have an affect here as well
 - MadGraph overpredicts and Sherpa underpredicts

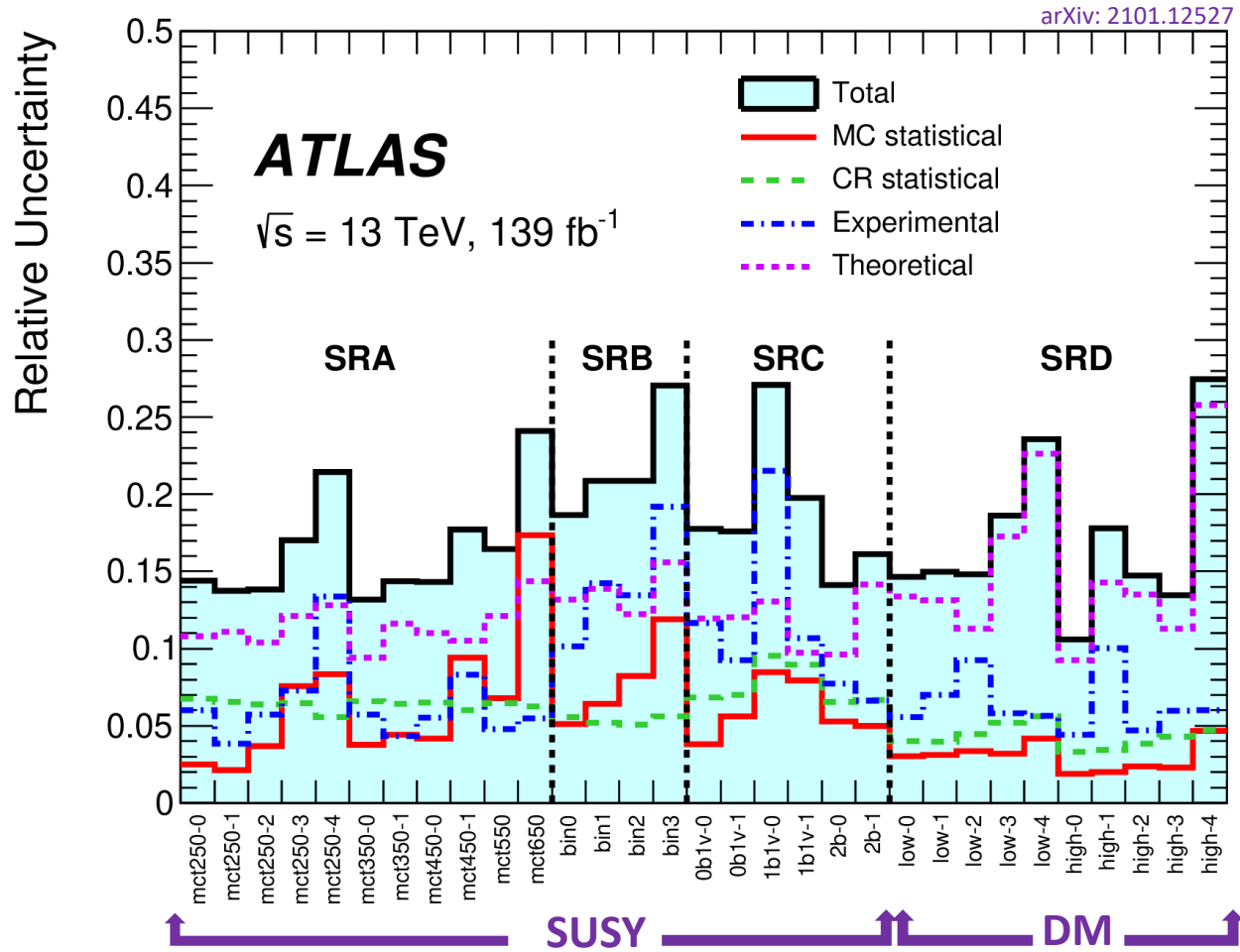


N.B: Overall experiment and theory systematic uncertainties of $\sim 13\%$ not shown in plots above

Data and SR from arXiv:2101.12527 shown here in comparison to SM theory predictions

- The below plots show the $\cos(\theta_{bb})^*$ distributions with Z+jets fitted from CR constraints for high and low mass mediators extrapolated to the SR
- The central value for Z+jets is taken from Sherpa but a generator systematic has been included (in the shaded region) to account for the difference in prediction from MadGraph
- The last two bins in both distributions were blinded as this is where we have the most discriminating power but is also where we saw the most disagreement between Sherpa and MadGraph





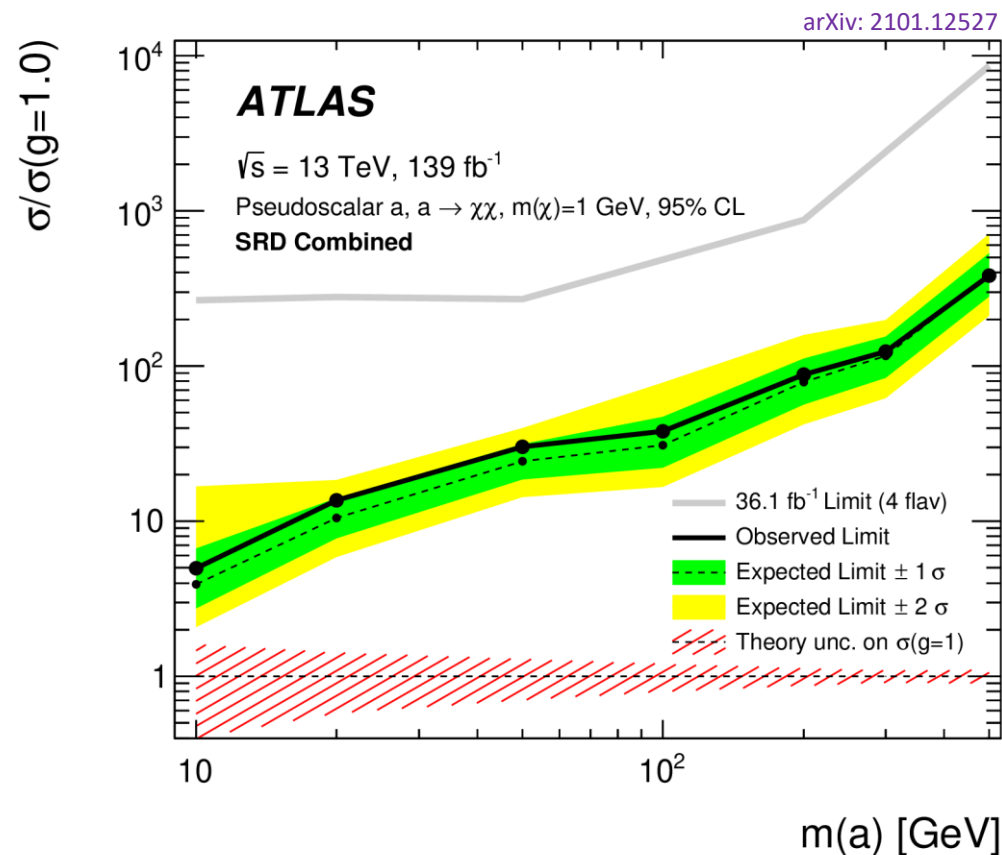
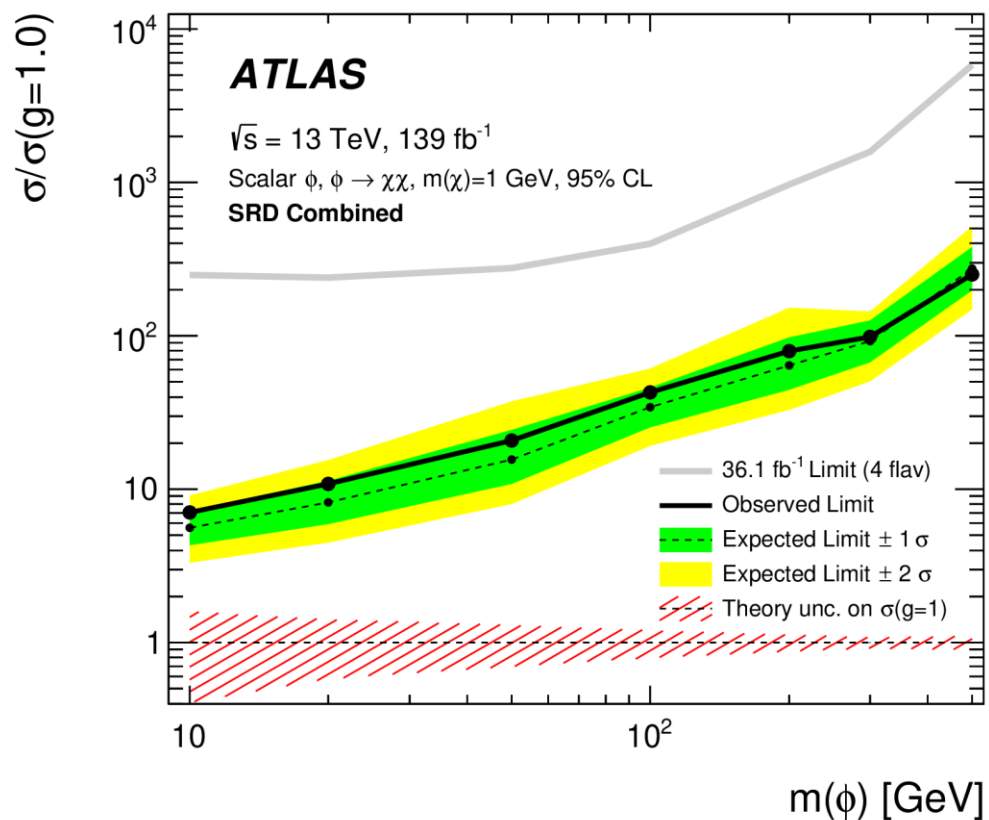
Post-fit relative systematic uncertainties

➤ CR constraints already taken into account

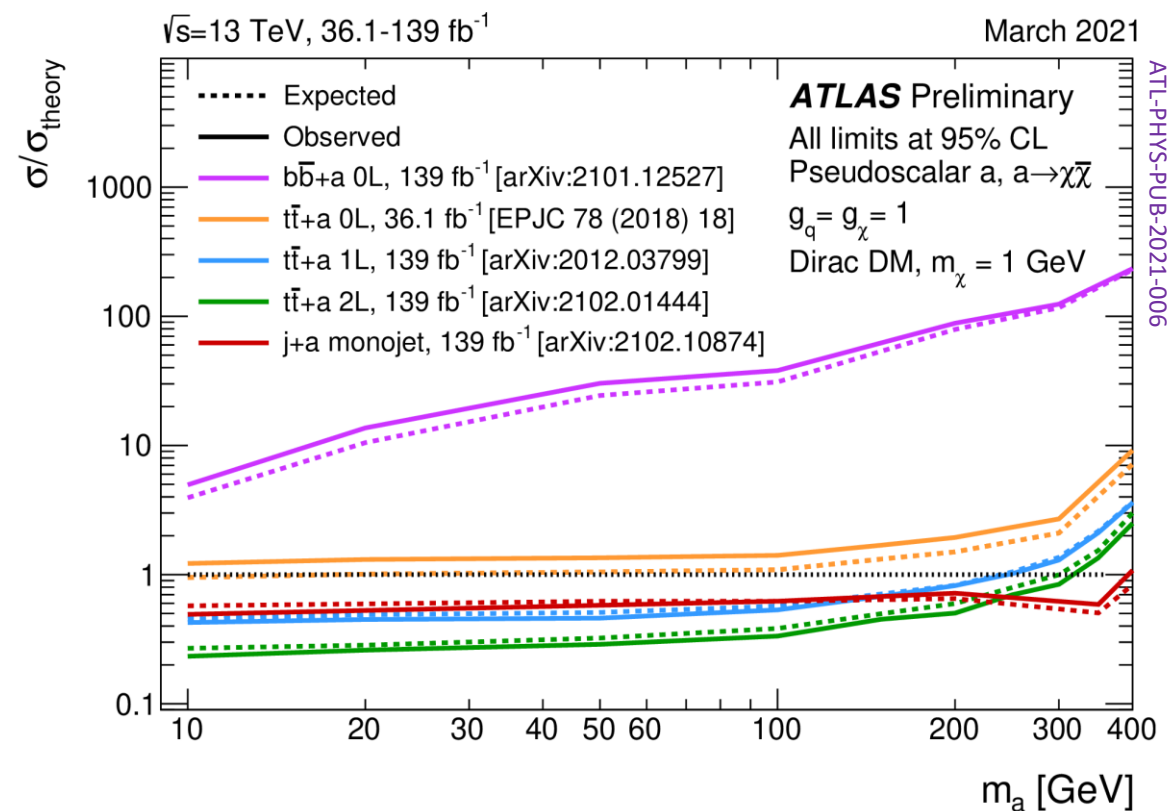
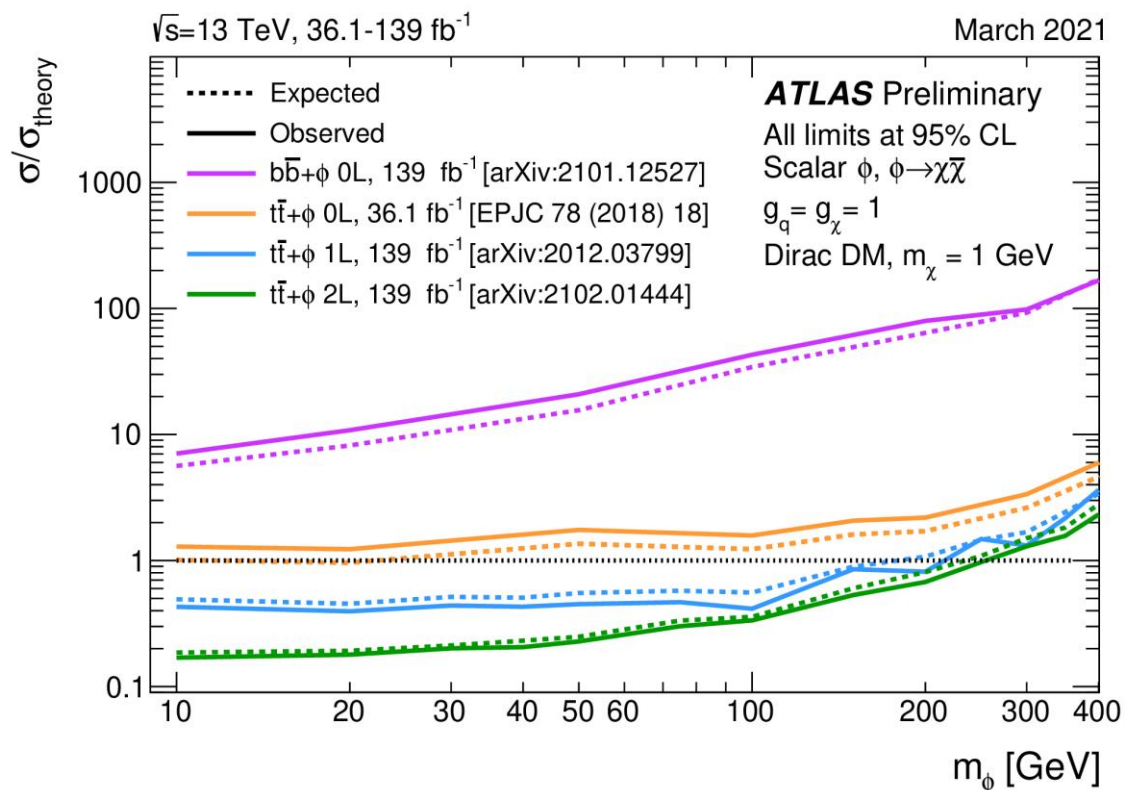
Majority of regions dominated by theoretical uncertainties

➤ Especially DM

- New search results around 10x more sensitive than previous 36.1fb^{-1} results but no evidence of DM
- small excess (<1 sigma) in high $\cos(\text{bb})^*$ weakens limits compared to expected
 - generator selection has a large affect on conclusion – understanding of Z and theory uncertainties could be the difference between a discovery and not



➤ New summary plot for 2021 for s-channel mediators



ATL-PHYS-PUB-2021-006

- Search for third generation SUSY and DM with no evidence for either found
- Limits on DM produced in association with b-jets improved by about 10x over previous iteration
- Theory uncertainties on modeling of SM backgrounds was a limiting factor in this analysis
- DM limits or deviation from SM could be stronger with better Monte Carlo prediction of the background in this phase space
- Further measurements of these SM backgrounds to searches in these regions of phase space needed

Backup

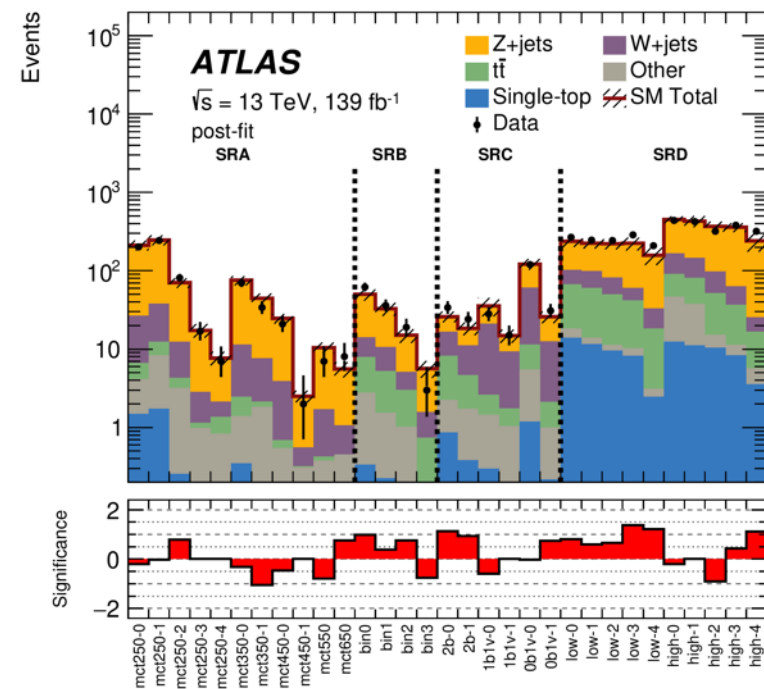
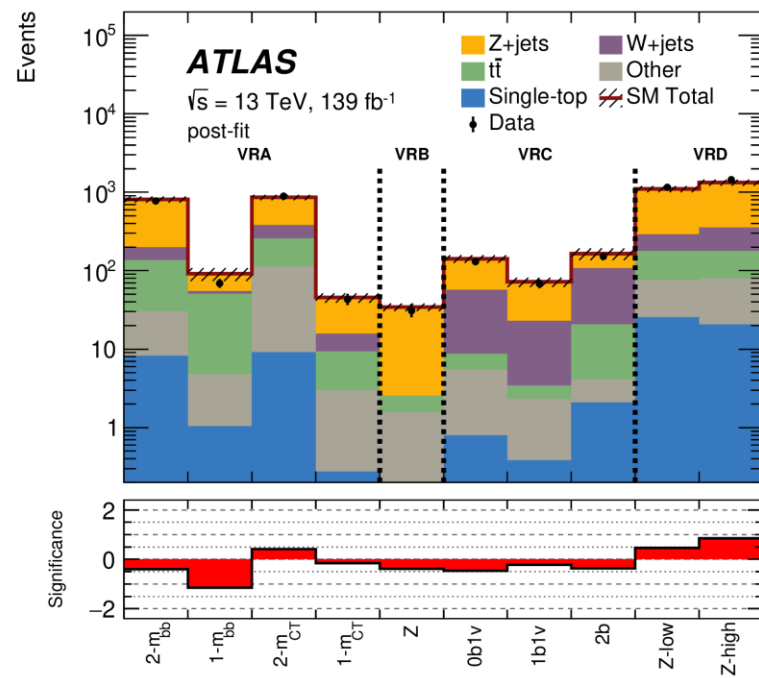
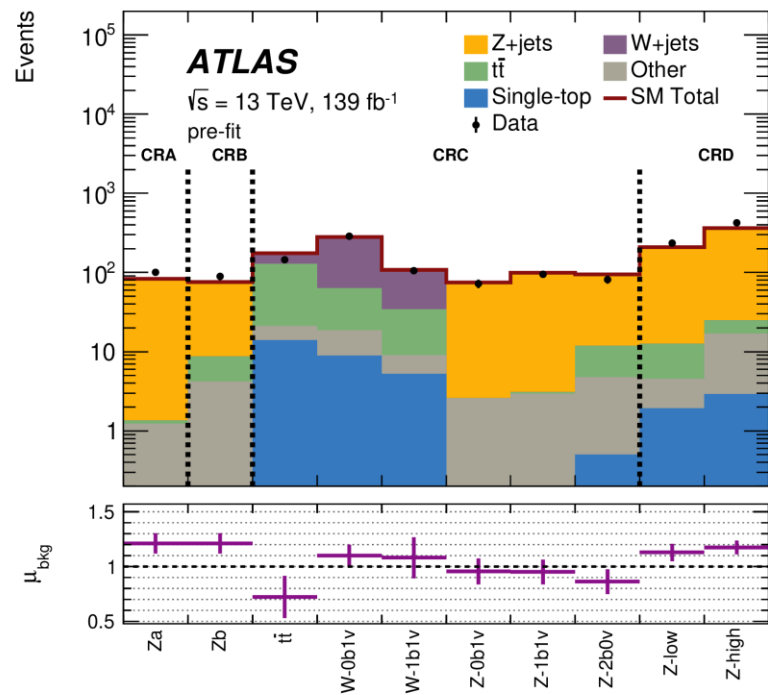
Variable		SRA	CRzA	$VR_{A1}^{m_{CT}}$	$VR_{A1}^{m_{bb}}$	$VR_{A2}^{m_{CT}}$	$VR_{A2}^{m_{bb}}$
Number of baseline leptons		0	2			0	
Number of high-purity leptons		–	2 SFOS			–	
$p_T(\ell_1)$	[GeV]	–	> 27			–	
$p_T(\ell_2)$	[GeV]	–	> 20			–	
$m_T(\ell, \mathbf{p}_T^{\text{miss}})$	[GeV]	–	> 20			–	
$m_{\ell\ell}$	[GeV]	–	[81, 101]			–	
Number of jets						$\in [2, 4]$	
Number of b -tagged jets						2	
j_1 and j_2 b -tagged						✓	
$p_T(j_1)$	[GeV]					> 150	
$p_T(j_2)$	[GeV]					> 50	
$p_T(j_4)$	[GeV]					< 50	
$\min[\Delta\phi(\text{jet}_{1-4}, \mathbf{p}_T^{\text{miss}})]$	[rad]					> 0.4	
E_T^{miss}	[GeV]	> 250	< 100			> 250	
$\tilde{E}_T^{\text{miss}}$	[GeV]	–	> 250			–	
$E_T^{\text{miss}}/m_{\text{eff}}$		> 0.25	–			–	
$\tilde{E}_T^{\text{miss}}/m_{\text{eff}}$		–	> 0.25			–	
m_{bb}	[GeV]	> 200		< 200	> 200	< 200	> 200
m_{CT}	[GeV]	> 250		> 250	[150, 250]	> 250	[150, 250]
m_{eff}	[GeV]	> 500			[500, 1500]		> 1500

Variable		SRB	CRzB	VRzB
Number of baseline leptons		0	2	
Number of high-purity leptons		–	2 SFOS	
$p_T(\ell_1)$	[GeV]	–	> 27	
$p_T(\ell_2)$	[GeV]	–	> 20	
$m_{\ell\ell}$	[GeV]	–	[76, 106]	
$m_T(\ell, \mathbf{p}_T^{\text{miss}})$	[GeV]	–	> 20	
Number of jets			$\in [2, 4]$	
Number of b -tagged jets			2	
$p_T(j_1)$	[GeV]		> 100	
$p_T(j_2)$	[GeV]		> 50	
$\min[\Delta\phi(\text{jet}_{1-4}, \mathbf{p}_T^{\text{miss}})]$	[rad]		> 0.4	
j_1 not b -tagged		–	✓	–
E_T^{miss}	[GeV]	> 250	< 100	
$\tilde{E}_T^{\text{miss}}$	[GeV]	–	> 250	
m_{CT}	[GeV]		< 250	
w_{XGB}		> 0.85	[0.3, 0.63]	> 0.63

Variable		SRC-2b	SRC-1b1v	SRC-0b1v	VRC-2b	VRC-1b1v	VRC-0b1v
Number of jets					$\in [2, 5]$		
j_1 not b -tagged					✓		
Number of baseline leptons					0		
Number of b -tagged jets		≥ 2	1	0	≥ 2	1	0
N_{vtx}		≥ 0	≥ 1	≥ 1	≥ 0	≥ 1	≥ 1
m_{vtx}	[GeV]	–	> 0.6	> 1.5	–	> 0.6	> 1.5
$p_{\text{T}}^{\text{vtx}}$	[GeV]	–	> 3	> 5	–	> 3	> 5
$p_{\text{T}}(j_1)$	[GeV]	> 500	> 400	> 400	< 500	> 400	> 400
$E_{\text{T}}^{\text{miss}}$	[GeV]	> 500	> 400	> 400	< 500	> 400	> 400
$H_{\text{T};3}$	[GeV]	–	< 80	< 80	–	< 80	< 80
\mathcal{A}		> 0.80	> 0.86	–	[0.8, 0.9]	> 0.86	–
m_{jj}	[GeV]	> 250	> 250	–	[150, 250]	> 250	–
$\Delta\phi(j_1, b_1)$	[rad]	–	> 2.2	–	–	< 2.2	–
$\Delta\phi(j_1, \text{vtx})$	[rad]	–	–	> 2.2	–	–	< 2.2
$ \eta_{\text{vtx}} $		–	< 1.2	< 1.2	–	> 1.2	> 1.2

Variable		CRtC	CRwC-1b1v	CRwC-0b1v	CRzC-2b	CRzC-1b1v	CRzC-0b1v
j_1 not b -tagged					✓		
Number of high-purity leptons			1			2 SFOS	
$H_{T,3}$	[GeV]				< 80		
$p_T(j_1)$	[GeV]		> 400		> 300		> 400
$m_T(\ell, \mathbf{p}_T^{\text{miss}})$	[GeV]		[20, 120]				–
$m_{\ell\ell}$	[GeV]		–			[81, 101]	
E_T^{miss}	[GeV]		> 400			< 100	
$\tilde{E}_T^{\text{miss}}$	[GeV]		–		> 250		> 400
\mathcal{A}		> 0.5	> 0.8	–	> 0.5	> 0.8	–
m_{jj}	[GeV]	> 250	> 250	–	–	> 250	–
$N_{b\text{-jets}}$		≥ 2	1	0	≥ 2	1	0
N_{vtx}		–	≥ 1	≥ 1	–	≥ 1	≥ 1
m_{vtx}	[GeV]	–	> 0.6	> 1.5	–	> 0.6	> 1.5
p_T^{vtx}	[GeV]	–	> 3	> 5	–	> 3	> 5

Variable		SRD-low	SRD-high	CRzD-low	CRzD-high	VRzD-low	VRzD-high
Trigger plateau		$(p_T(j_1) - 20 \text{ GeV})(E_T^{\text{miss}} - 160 \text{ GeV}) > 5000 \text{ GeV}^2$					
N_{jets}		2–3					
$N_{b\text{-jets}}$		≥ 2					
$p_T(j_1)$	[GeV]	> 100					
$p_T(j_2)$	[GeV]	> 50					
$\min[\Delta\phi(\text{jet}_{1-3}, \mathbf{p}_T^{\text{miss}})]$	[rad]	> 0.4					
S		> 7					
$p_T(j_1)/H_T$		> 0.7					
Number of baseline leptons		0		2			0
Number of high-purity leptons		–		2 SFOS			–
$p_T(\ell_1)$	[GeV]	–		> 27			–
$p_T(\ell_2)$	[GeV]	–		> 20			–
$m_T(\ell, \mathbf{p}_T^{\text{miss}})$	[GeV]	–		> 20			–
$m_{\ell\ell}$	[GeV]	–		[81, 101]			–
$\tilde{E}_T^{\text{miss}}$	[GeV]	–		> 180			–
E_T^{miss}	[GeV]	> 180		< 100			> 180
$w_{D\text{-low}}^{tt}$		> 0	–		–	> 0	–
$w_{D\text{-low}}^Z$		> 0	–	> 0	–	[–0.2, 0]	–
$w_{D\text{-low}}^W$		> 0	–		–	> 0	–
$w_{D\text{-high}}^{tt}$		–	> 0		–	–	> 0
$w_{D\text{-high}}^Z$		–	> -0.1	–	> -0.1	–	[–0.3, –0.1]
$w_{D\text{-high}}^W$		–	> -0.05		–	–	> -0.05



SRD-high-SRD-low-0		SRD-low-1		SRD-low-2		SRD-low-3		SRD-low-4	
Z_theory_renorm	6.1 %	ttbar_theory_PS	6.0 %	Z_theory_renorm	7.5 %	Z_theory_GEN	11.3 %	Z_theory_GEN	14.2 %
Z_theory_ckkw_max	5.8 %	Z_theory_renorm	5.8 %	ttbar_theory_PS	5.3 %	Z_theory_renorm	9.8 %	Z_theory_qsf_max	10.2 %
ttbar_theory_PS	5.3 %	Z_theory_ckkw_max	4.7 %	JET_GroupedNP_1	5.3 %	mu_Z	5.2 %	Z_theory_renorm	8.5 %
Z_theory_qsf_max	5.0 %	Z_theory_GEN	4.6 %	mu_Z	4.5 %	ttbar_theory_PS	4.3 %	Z_theory_ckkw_max	8.4 %
st_theory_DS	4.1 %	mu_Z	4.0 %	JET_GroupedNP_2	4.2 %	Z_theory_qsf_max	4.2 %	ttbar_theory_GEN	5.8 %
mu_Z	4.0 %	Z_theory_qsf_max	3.8 %	JET_JER_EffectiveNP_1	4.1 %	ttbar_theory_GEN	3.7 %	mu_Z	5.6 %
ttbar_theory_GEN	3.7 %	ttbar_theory_GEN	3.7 %	Z_theory_ckkw_max	3.6 %	st_theory_DS	3.2 %	ttbar_theory_PS	4.7 %
W_theory_renorm	3.5 %	JET_GroupedNP_1	3.7 %	W_theory_renorm	3.0 %	JET_GroupedNP_1	2.8 %	JET_Flavor_Response	2.7 %
JET_GroupedNP_2	3.2 %	W_theory_renorm	3.3 %	Z_theory_qsf_max	2.6 %	Z_theory_ckkw_max	2.7 %	JET_JER_EffectiveNP_2	2.5 %
W_theory_fac	2.8 %	W_theory_fac	3.3 %	st_theory_DS	2.4 %	JET_JER_EffectiveNP_4	2.6 %	JET_JER_EffectiveNP_5	1.9 %
SRD-high-0		SRD-high-1		SRD-high-2		SRD-high-3		SRD-high-4	
Z_theory_renorm	5.9 %	Z_theory_GEN	10.9 %	Z_theory_renorm	9.1 %	Z_theory_renorm	10.1 %	Z_theory_qsf_max	15.8 %
Z_theory_GEN	4.0 %	Z_theory_renorm	6.6 %	Z_theory_GEN	7.4 %	mu_Z	4.3 %	Z_theory_ckkw_max	12.8 %
W_theory_renorm	3.7 %	JET_JER_EffectiveNP_6	5.7 %	mu_Z	3.8 %	Z_theory_qsf_max	3.3 %	Z_theory_GEN	11.5 %
mu_Z	3.3 %	JET_JER_EffectiveNP_5	5.1 %	Z_theory_ckkw_max	3.6 %	JET_GroupedNP_1	2.6 %	Z_theory_renorm	10.3 %
W_theory_fac	2.7 %	W_theory_renorm	3.7 %	W_theory_renorm	3.1 %	JET_JER_EffectiveNP_3	2.2 %	stat_4	5.4 %
JET_GroupedNP_1	2.3 %	mu_Z	3.4 %	ttbar_theory_PS	3.0 %	ttbar_theory_PS	2.2 %	mu_Z	4.7 %
JET_GroupedNP_2	2.0 %	ttbar_theory_PS	3.0 %	JET_GroupedNP_1	2.4 %	JET_JER_EffectiveNP_4	2.0 %	JET_JER_EffectiveNP_4	3.8 %
Z_theory_ckkw_max	2.0 %	W_theory_fac	2.6 %	Z_theory_qsf_max	2.1 %	JET_JER_EffectiveNP_7restTerm	1.7 %	JET_JER_EffectiveNP_7restTerm	2.7 %
ttbar_theory_PS	1.9 %	JET_JER_EffectiveNP_2	2.5 %	JET_GroupedNP_2	2.0 %	JET_JER_EffectiveNP_6	1.7 %	W_theory_renorm	2.5 %
FT_EFF_C_systematics	1.8 %	JET_JER_EffectiveNP_7restTerm	2.4 %	W_theory_fac	1.8 %	st_theory_DS	1.7 %	JET_JER_EffectiveNP_3	1.9 %

