Searching for New Physics with ATLAS: Anomaly Detection in Multilepton Final States

Atanay Odella[1], on behalf of the ATLAS Collaboration

[1] Institut de Física d'Altes Energies, Barcelona

CPAN
Ingenio 2010

XVII CPAN DAYS

ATLAS EXPERIMENT

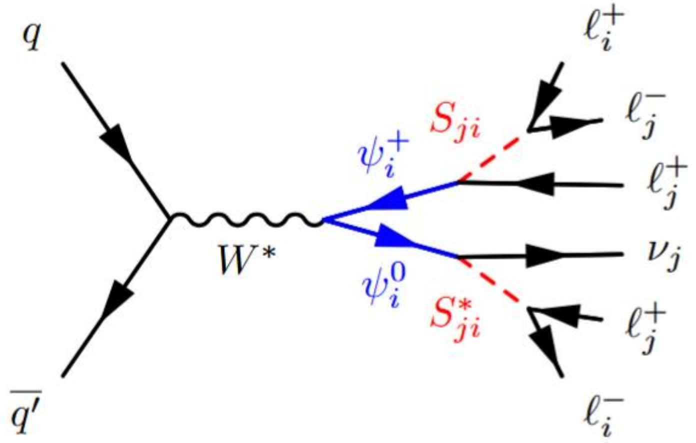
IFAE
Institut de Física d'Altes Energies

EXCELENCIA SEVERO OCHOA

Introduction

Analysis Overview

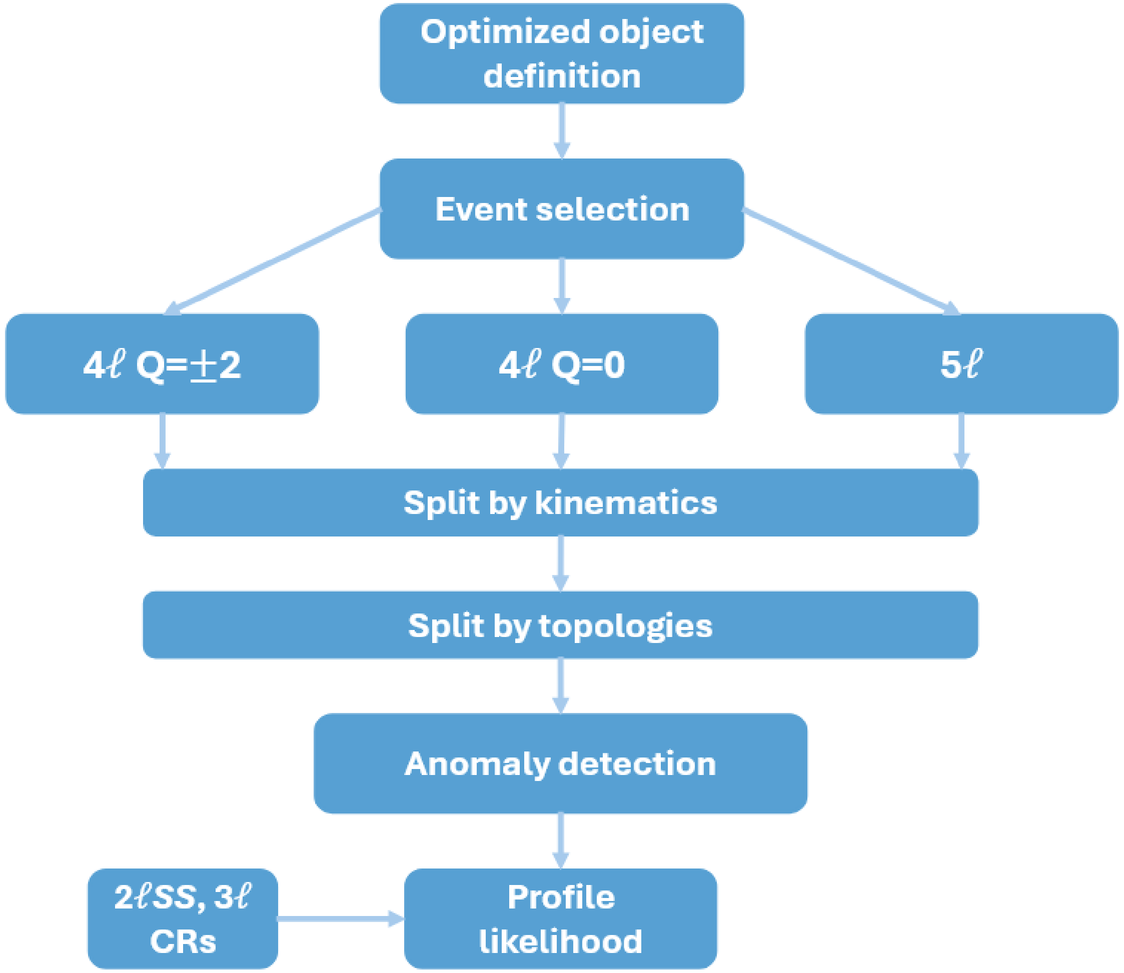
- The analysis targets **high lepton multiplicity** $(\geq 4\ell)$, making it well-suited for detecting heavy Beyond the Standard Model (BSM) resonances because of the low SM background.
- A new approach is implemented to exploit **charge** and **lepton flavour** across the **full Run 2** dataset, creating dedicated search regions where little-tested new physics may populate.
- A novel **anomaly detection** technique is used to boost our **sensitivity** to potential BSM signals in a model-agnostic way.
- The analysis is designed to target both **new physics** discovery, and sensitivity to a **broad range** of models.



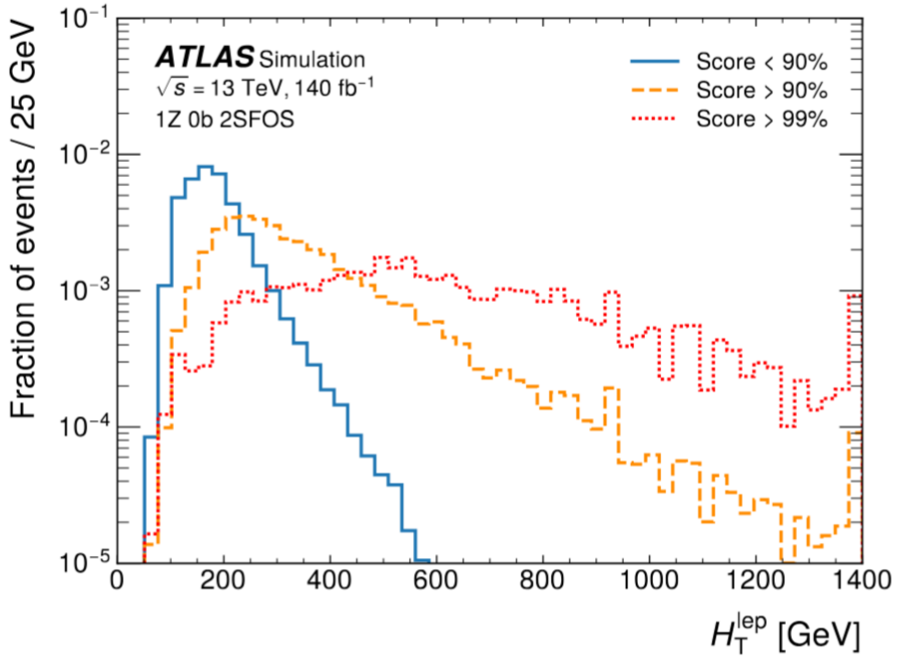
Vector-Like Lepton (VLL) decaying into a multilepton final state via scalars S_{ji}.

Analysis Strategy

- Acceptance maximised by using *Loose* leptons with $p_T > 10$ GeV.
- Phase space divided into distinct **topologies** based on lepton multiplicity and net charge to effectively separate significantly different backgrounds: $\mathbf{4\ell\ Q = 0}$, $\mathbf{4\ell\ Q = \pm 2}$, and $\geq \mathbf{5\ell}$.
- Each of these topologies exploited using **anomaly detection**.
- Resulting anomaly discriminant used in two optimised search strategies: **model-independent** and **model-dependent**.



- **Normalizing flows** used as density estimators to learn MC background probabilities, providing an anomaly score for **outlier detection**.
- Flows trained with physically motivated **high-level variables** (reduced set for $\geq 5\ell$ events) in each model-independent region.

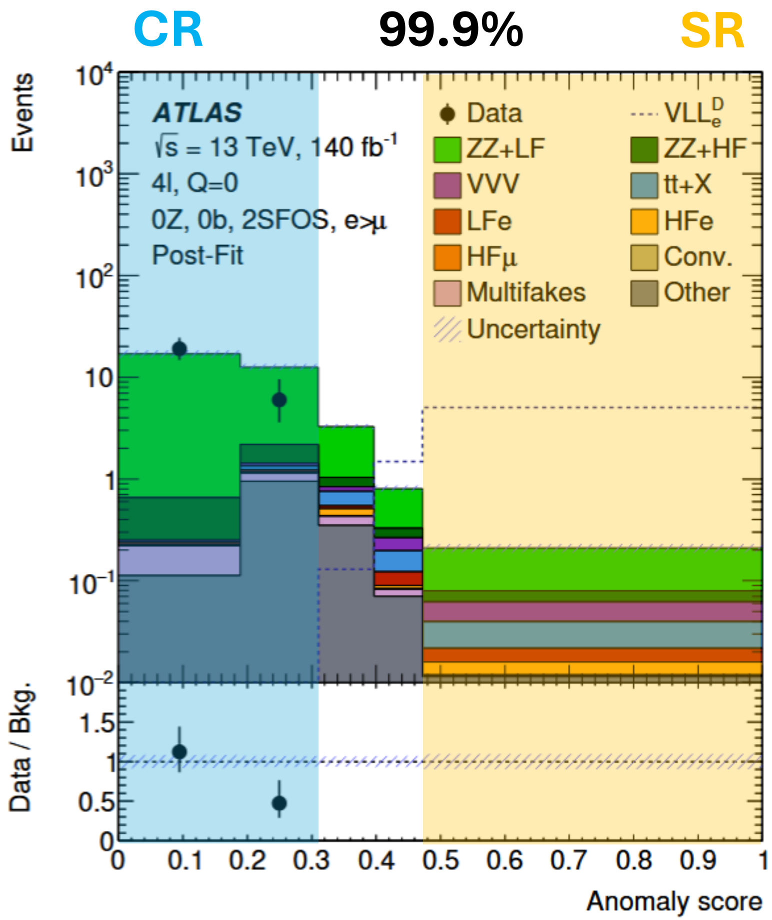


Training variable	Description
	Input to 4ℓ $Q = 0$, 4ℓ $Q = \pm 2$, and $\geq 5\ell$ 1Z
H_T^{lep}	Sum of transverse momenta of leptons
H_T^{jets}	Sum of transverse momenta of jets
E_T^{miss}	Missing transverse energy
N_{jets}	Number of jets
$p_T(Z)$	Transverse momentum of lepton pair closest to Z boson
	Input to 4ℓ $Q = 0$ and 4ℓ $Q = \pm 2$
$p_T(\ell\ell)$	Transverse momentum of lepton pair second-closest to Z boson
$m(Z)$	Invariant mass of lepton pair closest to Z boson
$m(\ell\ell)$	Invariant mass of lepton pair second-closest to Z boson
$m^{high}(3\ell)$	Largest invariant mass of lepton pair closest to Z boson and another lepton
$m^{low}(3\ell)$	Smallest invariant mass of lepton pair closest to Z boson and another lepton
$m(4\ell)$	Invariant mass of four-lepton system
$m_T(4\ell, E_T^{miss})$	Transverse mass of four leptons and E_T^{miss}
$m_T(Z, E_T^{miss})$	Transverse mass of lepton pair closest to Z boson and E_T^{miss}
$m_T(\ell\ell, E_T^{miss})$	Transverse mass of lepton pair second-closest to Z boson and E_T^{miss}
$\sum_{i=1}^{jets} \text{pcb}_i$	Sum of pseudo-continuous b-tagging score

Anomaly score regions

Model-independent strategy

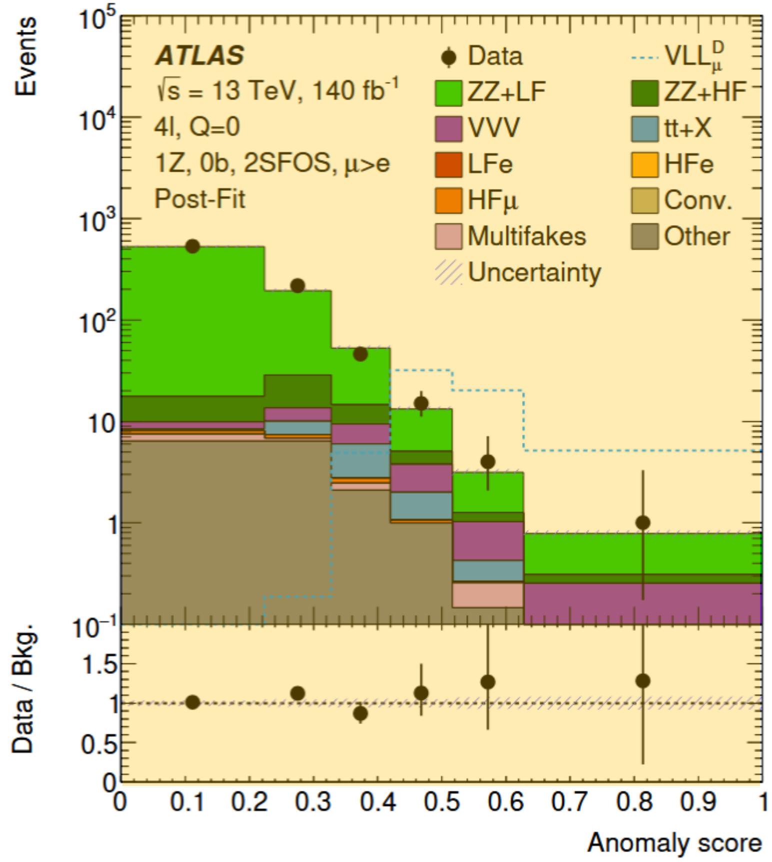
- **Model-independent** search regions optimised and separated by number of Z candidates and number of same-flavour-opposite-sign pairs (SFOS).
- **50%, 90%, 99%** and **99.9%** background rejection points defined as bins, ensuring each one has ≥ 0.1 background events and <20% MC Statistical error (otherwise dropped).
- **Control regions** (CRs) are defined as the <90% bins, whereas the >90% (when available) become the **signal regions** (SRs).
- Each discovery SR is **fitted** at a time with CRs and low-anomaly score regions to give model-independent **significances** and **limits**.



Model-dependent strategy

- **Model-dependent** search regions further separated by flavour and presence of b-jets.
- Highest anomaly score bin defined with 0.1 background and 20% MC statistical error. New bins made by increasing the background yield from right to left by a factor of 4 (until no longer possible).
- **Whole distribution** is used as a **signal region**.
- All regions are simultaneously fitted along with CRs to provide **limits** on benchmark models.

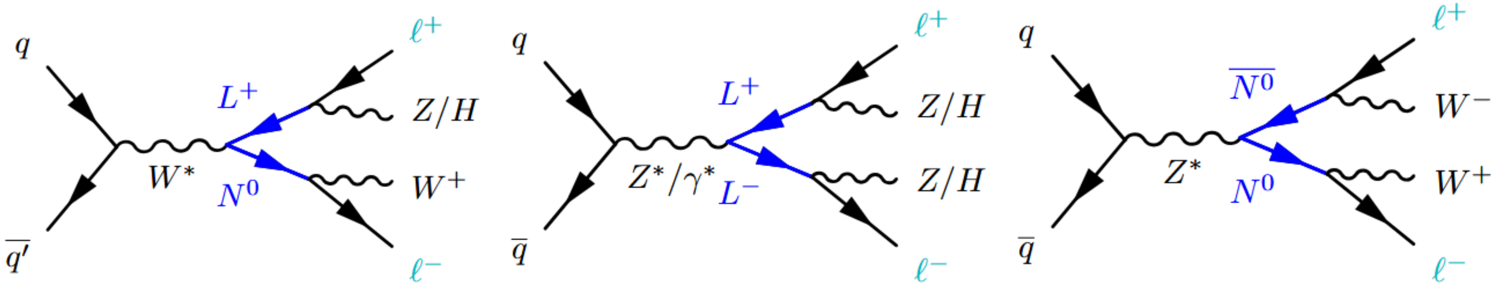
SR



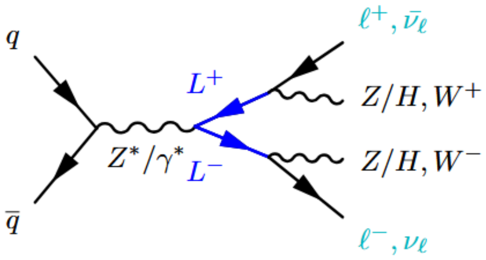
Signal benchmarks

Vector-Like Leptons (VLLs)

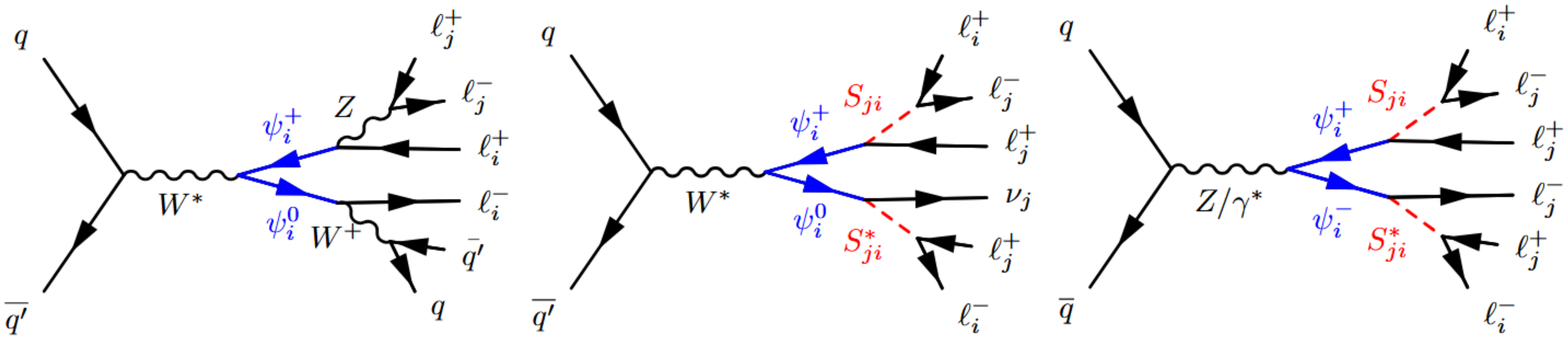
doublet



singlet

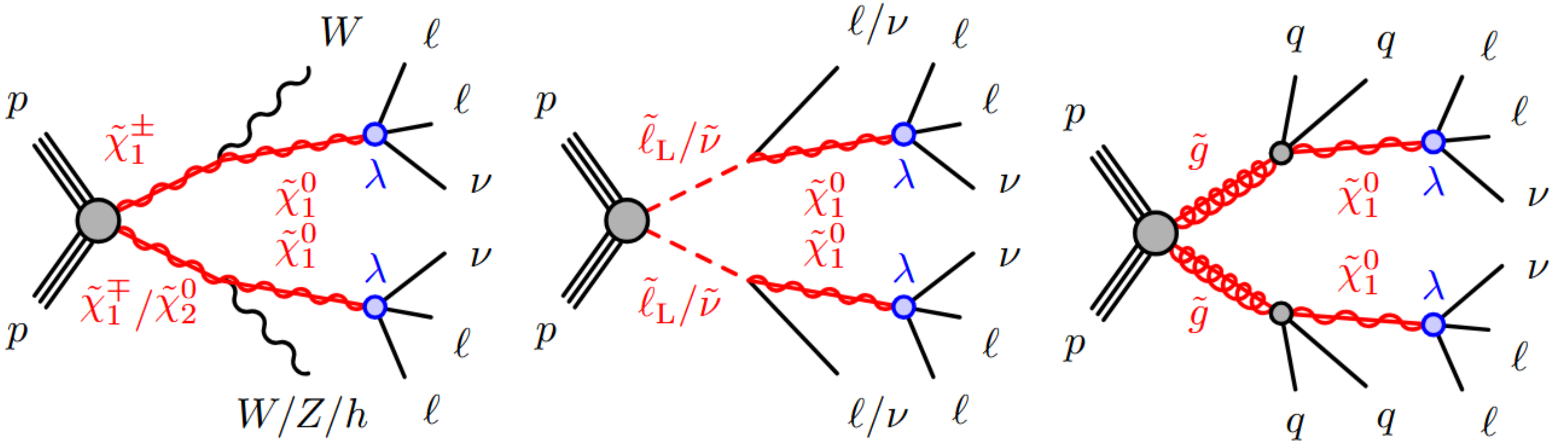


flavourful

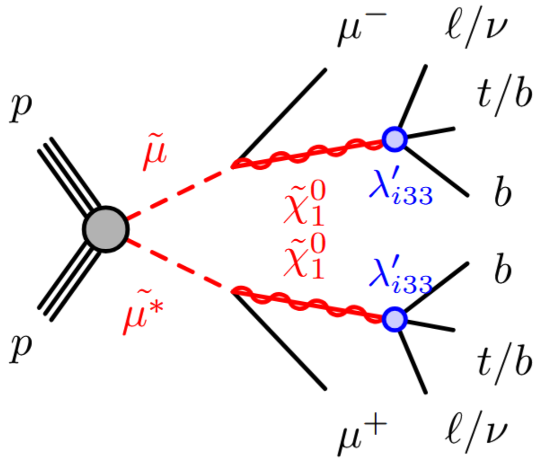


Signal benchmarks

Wino



Smuon



Anomaly score regions

Model-dependent

Topology	Splitting	Number of regions	Labelling	Example final states
4ℓ $Q = 0$	Number of Z candidates 0 or ≥ 1 b-jet SFOS Flavour of non-Z leptons	20	0Z 0b 0SFOS 0Z 0b 1SFOS $e > \mu$ 0Z 0b 1SFOS $e < \mu$ 0Z 0b 2SFOS $e > \mu$ 0Z 0b 2SFOS $e = \mu$ 0Z 0b 2SFOS $e < \mu$ 1Z 0b 1SFOS 1Z 0b 2SFOS $e > \mu$ 1Z 0b 2SFOS $e < \mu$ 2Z 0b (×2 with $\geq$ 1b)	$e^+e^+\mu^-\mu^-$ $e^+e^-e^+\mu^-$ $\mu^+\mu^-e^+\mu^-$ $e^+e^-e^+e^-$ $e^+e^-\mu^+\mu^-$ $\mu^+\mu^-\mu^+\mu^-$ $Z(\to \ell\ell), e^+\mu^-$ $Z(\to \ell\ell), e^+e^-$ $Z(\to \ell\ell), \mu^+\mu^-$ $ZZ(\to 4\ell)$
4ℓ $Q = \pm 2$	0 or ≥ 1 b-jets Flavour of non-Z leptons	6	0b $e > \mu$ 0b $e = \mu$ 0b $e < \mu$ (×2 with $\geq$ 1b)	$e^+e^+e^+e^-$ $e^+e^+\mu^+\mu^-$ $\mu^+\mu^+\mu^+\mu^-$
$\geq 5\ell$	Number of Z candidates 0 or ≥ 1 b-jets Flavour of non-Z leptons (0 b-jet)	6	5ℓ 0Z 0b $e > \mu$ 5ℓ 0Z 0b $e < \mu$ 5ℓ 0Z $\geq$ 1b 5ℓ $\geq$1Z 0b $e > \mu$ 5ℓ $\geq$1Z 0b $e < \mu$ 5ℓ $\geq$1Z $\geq$ 1b	$e^+e^-e^+e^-\mu^-$ $\mu^+\mu^-e^+\mu^-\mu^-$ $ZZ(\to 4\ell), e^+$ $Z(\to \ell\ell), \mu^+\mu^+e^-$

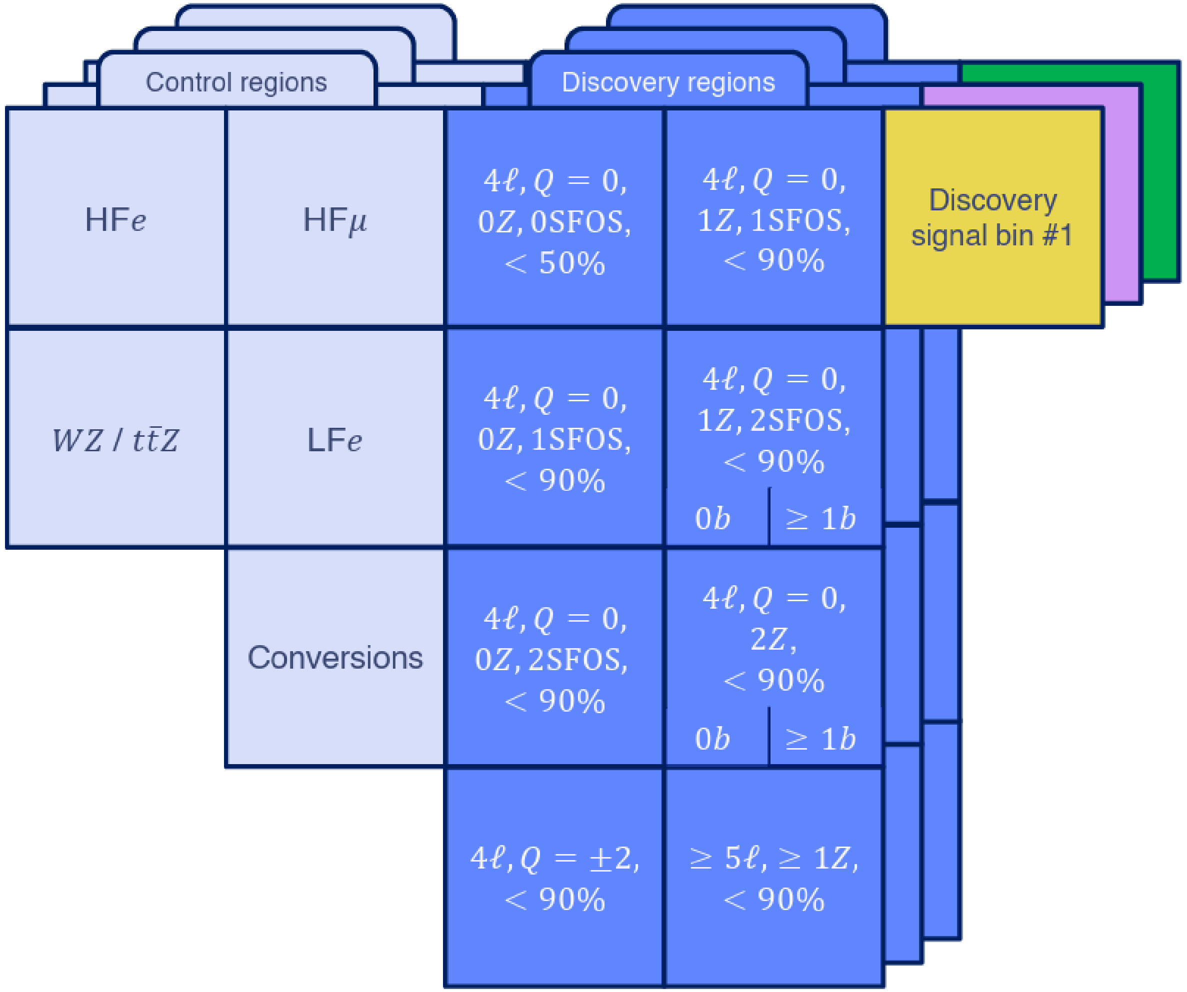
Model-independent

Topology	Splitting	Number of regions	Labelling	Example final states
4ℓ $Q = 0$	Number of Z candidates SFOS 0 or ≥ 1 b-jet (only low-anomaly score 1Z 2SFOS and 2Z regions)	6 (+2 low-anomaly score regions)	0Z 0SFOS 0Z 1SFOS 0Z 2SFOS 1Z 1SFOS 1Z 2SFOS 2Z	$e^+e^+\mu^-\mu^-$ $e^+e^-e^+\mu^-$ $e^+e^-\mu^+\mu^-$ $Z(\to \ell\ell), e^+\mu^-$ $Z(\to \ell\ell), \mu^+\mu^-$ $ZZ(\to 4\ell)$
4ℓ $Q = \pm 2$	-	1	$Q = \pm 2$ incl.	$e^+e^+\mu^+\mu^-$
$\geq 5\ell$	Number of Z candidates	2	5ℓ 0Z 5ℓ $\geq$1Z	$e^+e^-e^+\mu^-\mu^-$ $Z(\to \ell\ell), e^+e^-\mu^+$

- **High granularity** in model-dependent regions to maximise **sensitivity**, whereas **low granularity** in model-independent to reduce **look-elsewhere effect**.

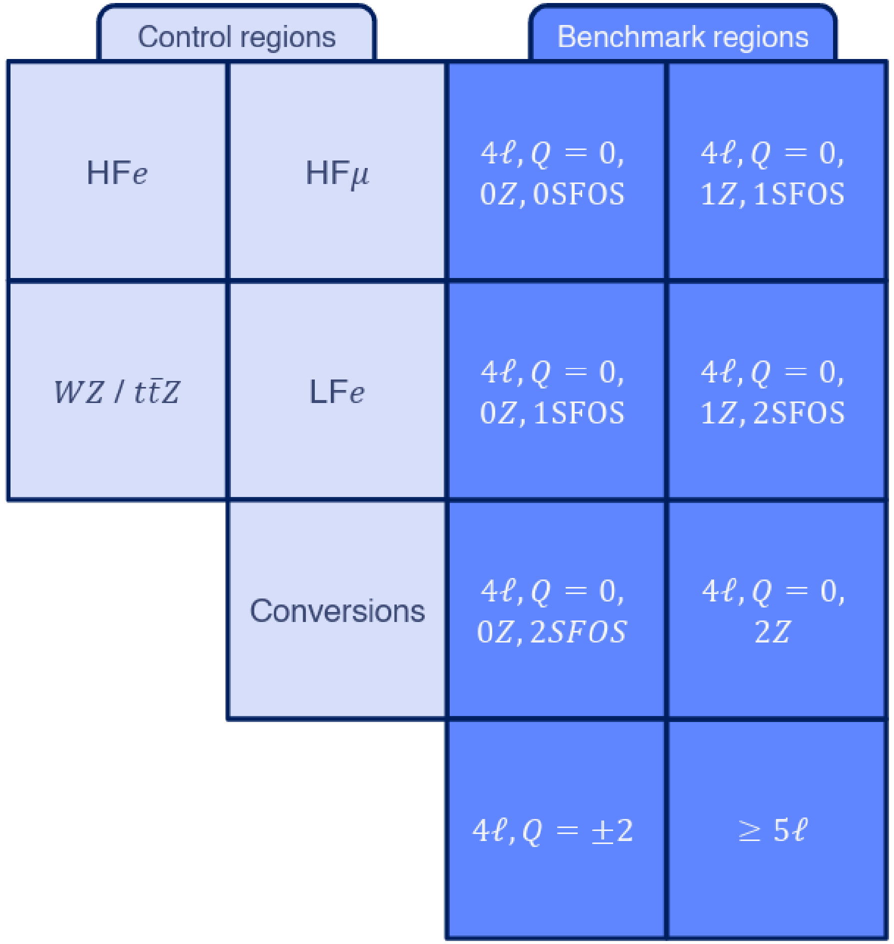
Strategy summary

Model-independent



Each layer fitted separately.

Model-dependent



Regions further split by non-Z lepton flavour and presence of b-jets.

Background corrections

Control region definition

- A **template fit** is employed to achieve a semi-data-driven estimation of the complex fake lepton background, which includes single and multi-fake groupings from **light flavour**, **heavy flavour**, and internal/material **conversions**.

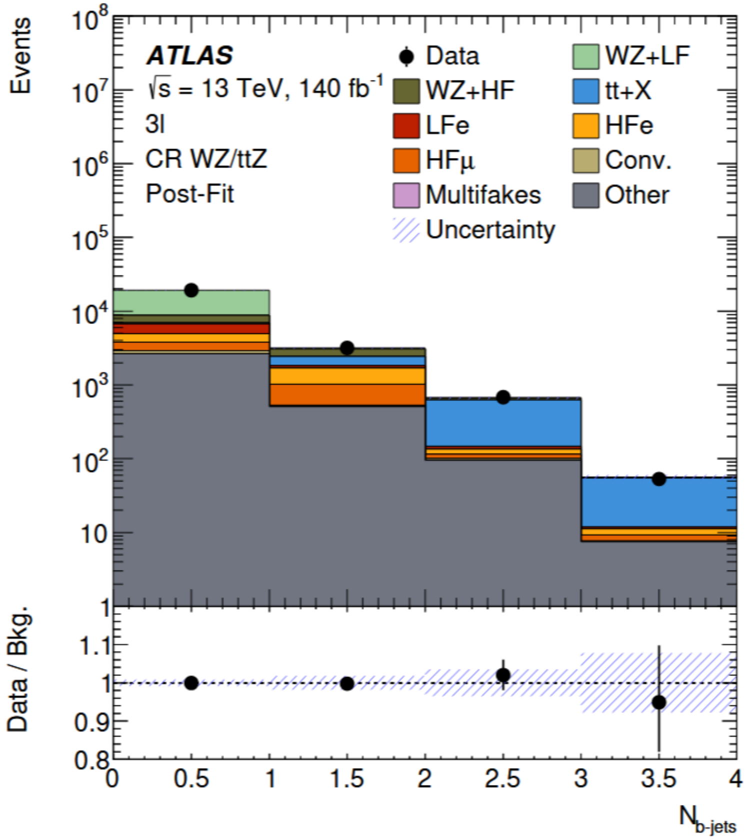
- **Control regions** in the 2ℓSS and 3ℓ topologies are used to fit **normalization factors** (NF) for each specific fake type. A dedicated 3ℓ CR fits NFs for WZ and $t\bar{t}Z$. These NFs are then applied **per-lepton** for events with multiple fake sources.

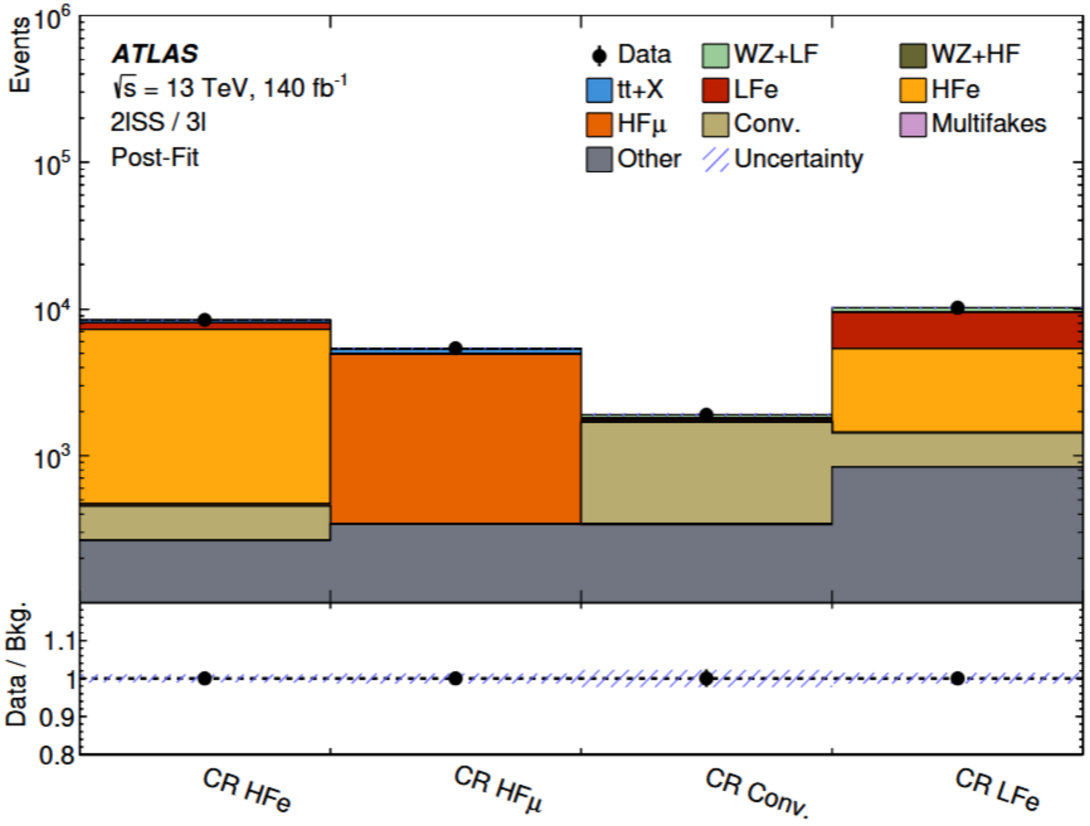
Table 4: Summary of the event selection applied in the $2\ell SS$ and 3ℓ template fit control regions.

	$2\ell SS$		3ℓ		
	HF_e	HF_μ	Conversion	LF_e	$W^\pm Z/\ t\bar{t}Z$
Lepton flavour	μe	$\mu\mu$	$e\mu\mu \parallel \mu e\mu \parallel \mu\mu e$		any
$\ell\ p_\mathrm{T}$ [GeV]	$> (20, 10)$		$> (10, 20, 20)$	$> 20(10)\ \mu\ (e)$	$> (10, 20, 20)$
Total charge	± 2		± 1		
N_jets	≥ 2		-	-	≥ 1
$N_{b-\mathrm{jets}}$	1 @ 77%		0 @ 77%	0 @ 85%	≥ 0 @ 77%
Z candidates	No		0 Z	1 Z $(\mu\mu)$	1 Z $(ee \parallel \mu\mu)$
Inv. Mass [GeV]	-		$\lvert M_{3\ell} - M_Z \rvert < 10$	-	-
Conv. candidate	Veto		Accept	Veto	
Prompt lepton BDT	Yes (leading)		-	-	-
Additional Cuts	QMisID BDT	-	-	$E_\mathrm{T}^\mathrm{miss} < 20$ GeV	-
	(sub-leading electron)	-	-		-

Control region definition

- The 2ℓSS and 3ℓ control regions are then used for fitting as explained with the model-independent and model-dependent regions.





Other background corrections

Data-driven QmisID estimate

- Events featuring a mis-identified charge (QmisID) lepton make a significant contribution to the Q=$\pm$2 regions. The QmisID rate is derived from 2ℓSS $Z \to \ell\ell$ events and subsequently applied to backgrounds in 4ℓ Q=0.

VV Njet reweighting

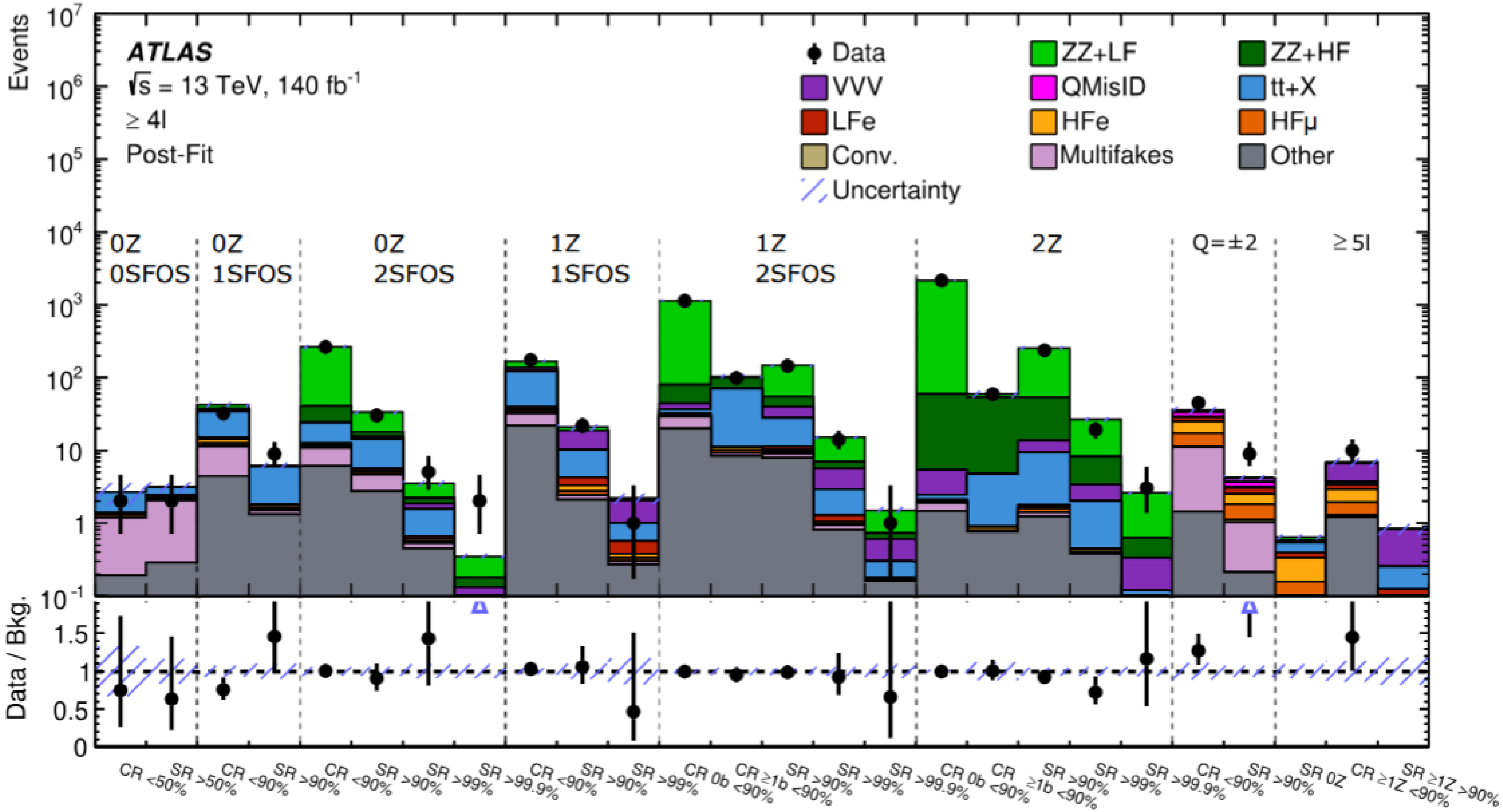
- Correction derived from 3ℓ control region targeting WZ+LF background.

tt NNLO reweighting

- The distributions for parton-level p_T, $m_{t\bar{t}}$, and $p_{T_{t\bar{t}}}$ are reweighted to match NNLO-QCD and NLO-EWK predictions. Correction is minor since the process only contributes as fakes to the SRs and as a background to the CRs.

Model-independent results

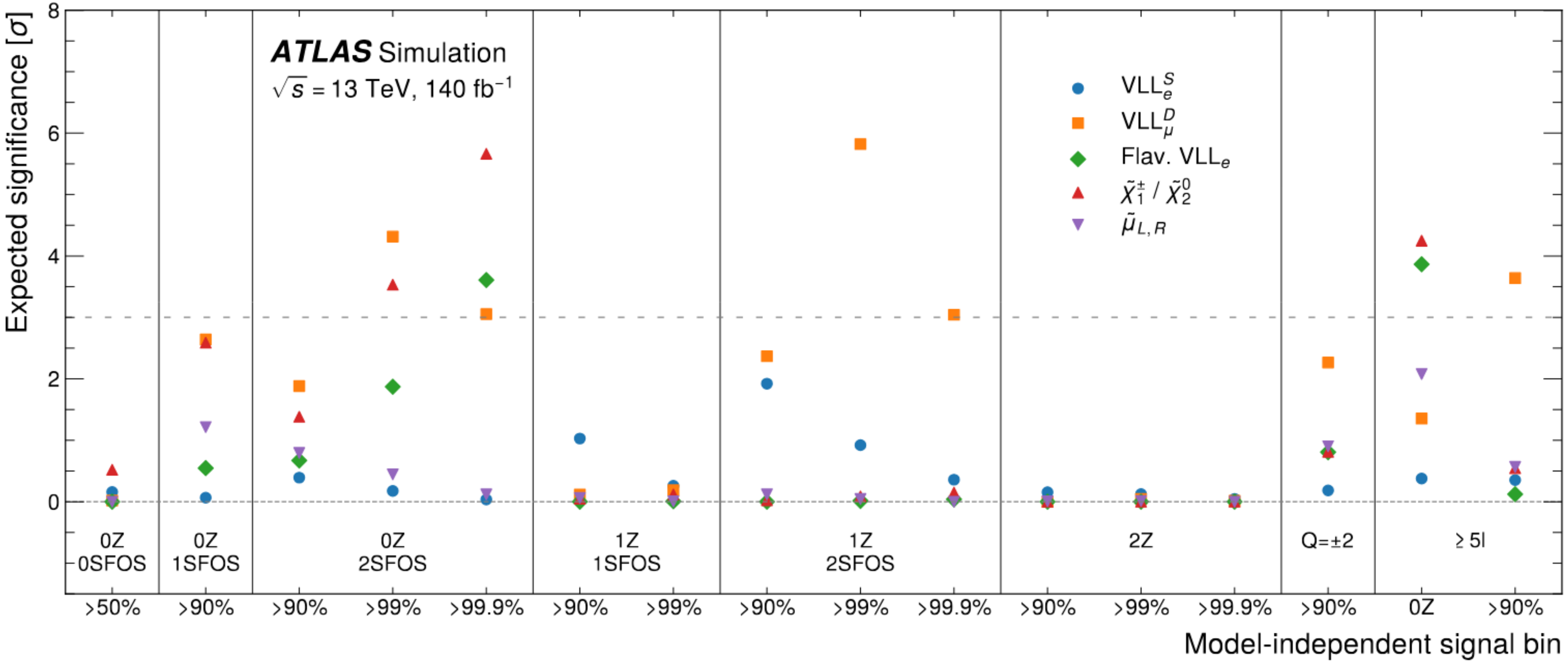
Model-independent post-fit regions



- Fitted model-independent anomaly score regions. Largest excesses in the **Q=0 0Z 2SFOS >99.9%** and **Q=±2 >90%** regions, with highest significance of **2.0σ local** and **0.89σ global**.

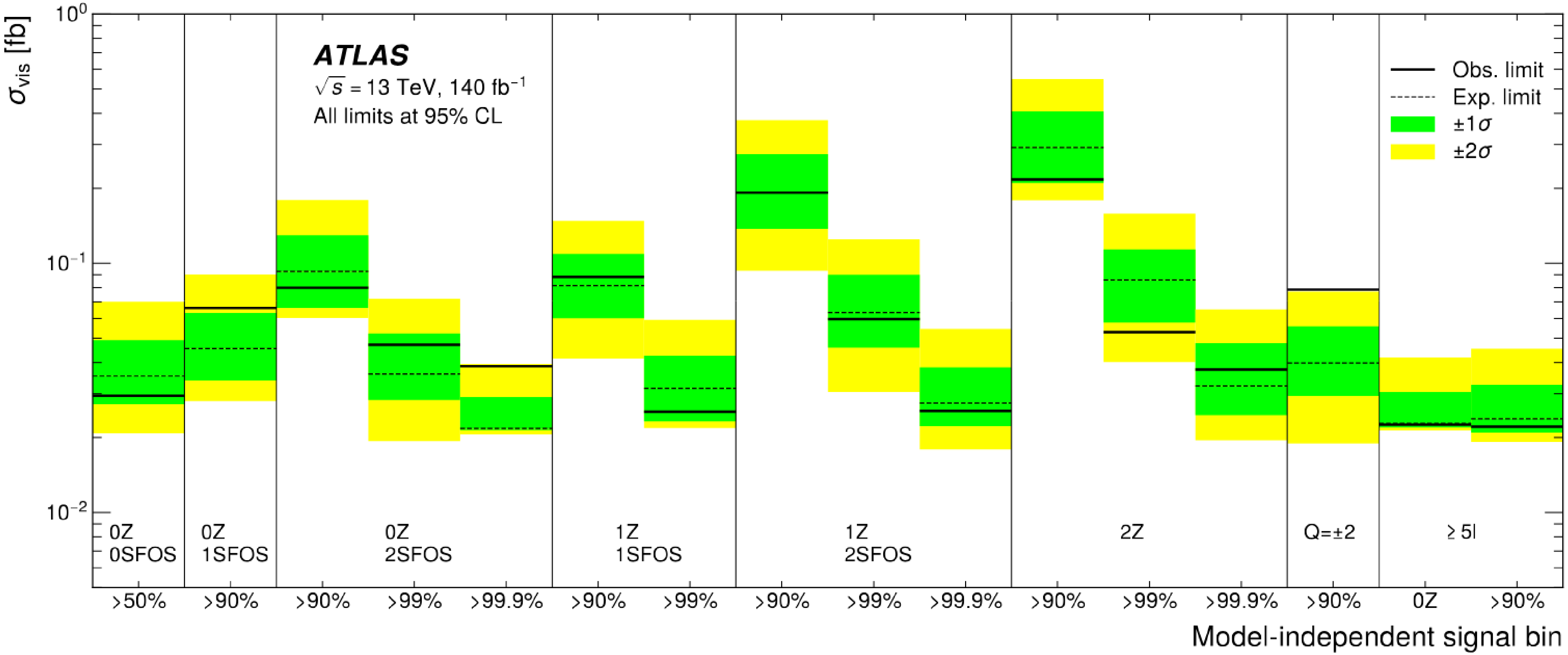
Model-independent discovery significance

- The **discovery significance** is reported for each model-independent signal bin. To demonstrate the useful sensitivity of these bins, the **signal benchmaks** are injected, and the expected significance is plotted using each bin individually (several SRs exceeding 2σ).



Model-independent limits

- After performing the model-independent fit, the **95% CL limit** (using 5000 B-only and S+B toys) is computed for each signal region.

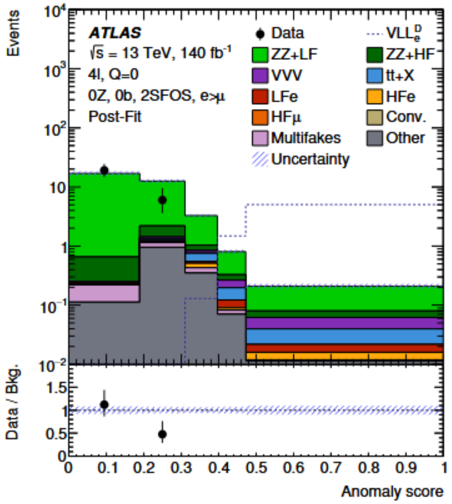


Model-dependent results

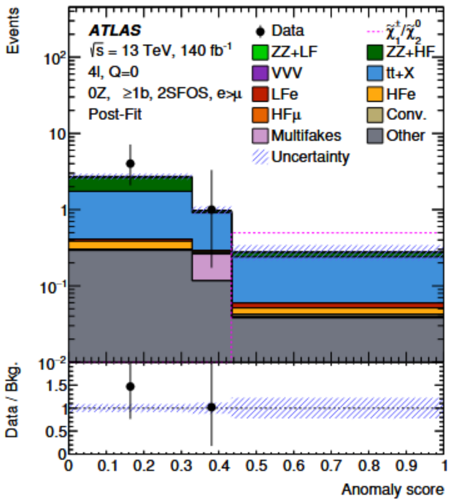
Model-dependent signal regions (subset)

- The anomaly score distribution is compared between data and the background estimate, following a background-only fit applied to the data in various benchmark regions.
- As anticipated, the signal predominantly populates the signal-like bins at **high anomaly scores**.
- Main backgrounds:
  - 4ℓ Q=0 (0b): ZZ
  - 4ℓ Q=0 ($\geq$ 1b): ttZ
  - 4ℓ Q=±2: QmisID + Wjets/ttbar fakes
  - $\geq$ 5ℓ: VVV + ZZ/ttbar fakes

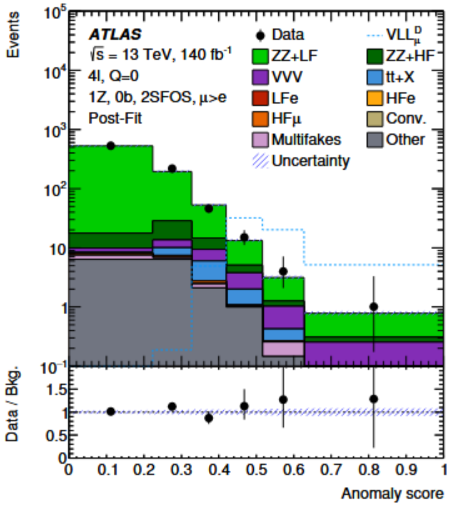
VLLe (doublet)



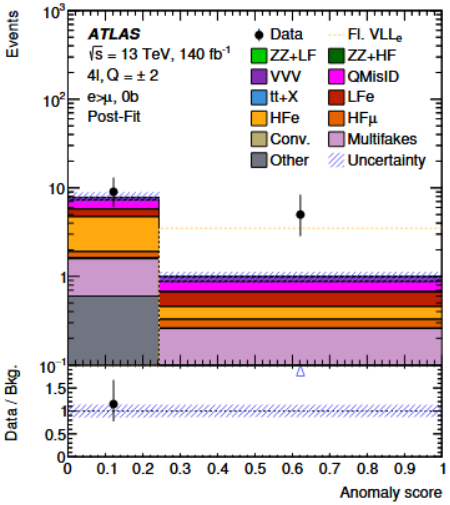
Wino



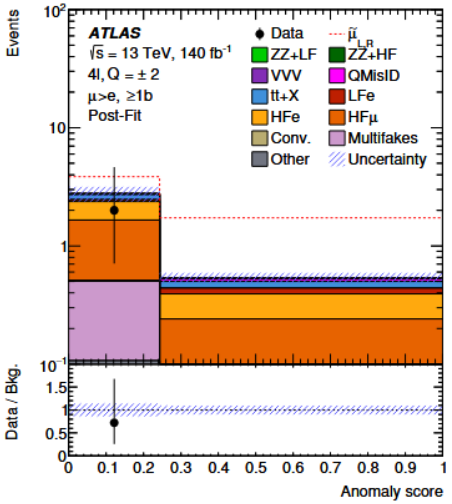
VLLμ (doublet)



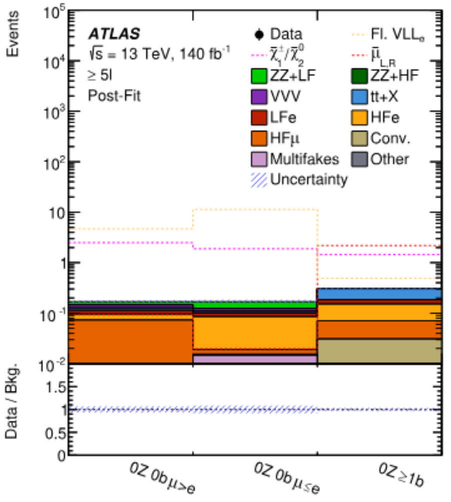
Fl. VLLe



Smuon



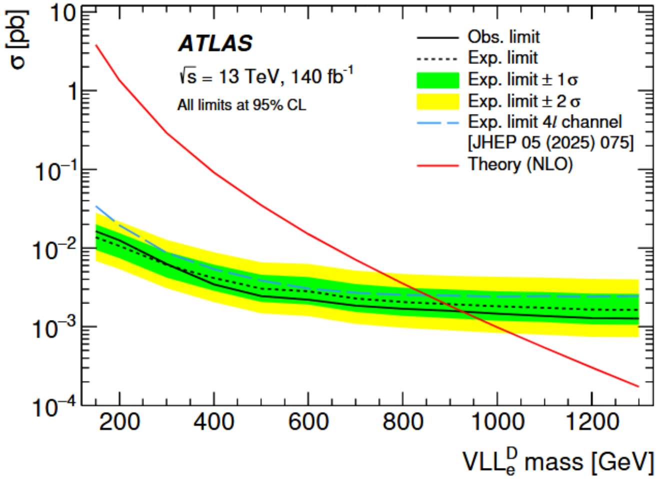
Fl. VLLe, Wino, Smuon



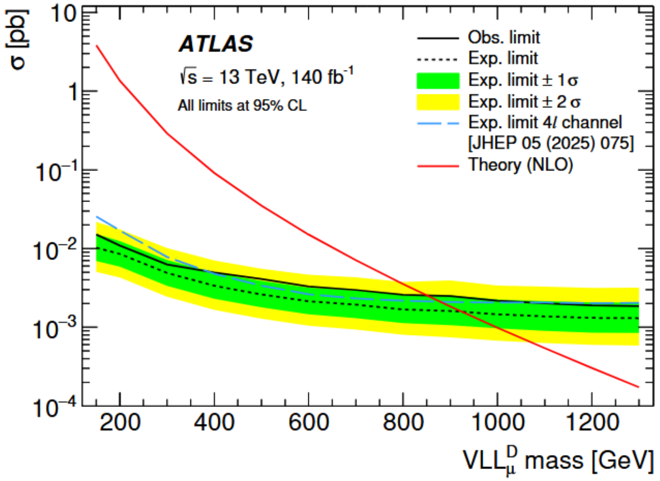
Model-dependent benchmark limits

- Limits are produced by performing the model-dependent fit across all regions simultaneously.
- **No significant excesses** are observed in any benchmark model tested.
- The resulting **sensitivity** is **competitive** with dedicated searches, demonstrating slightly superior performance in certain cases

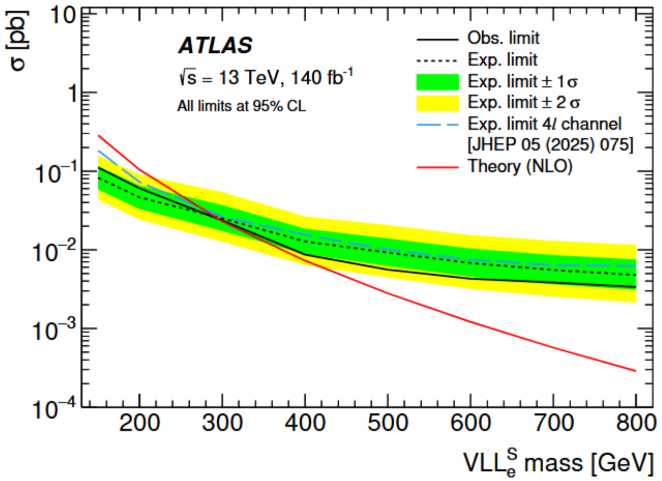
VLLe (doublet)



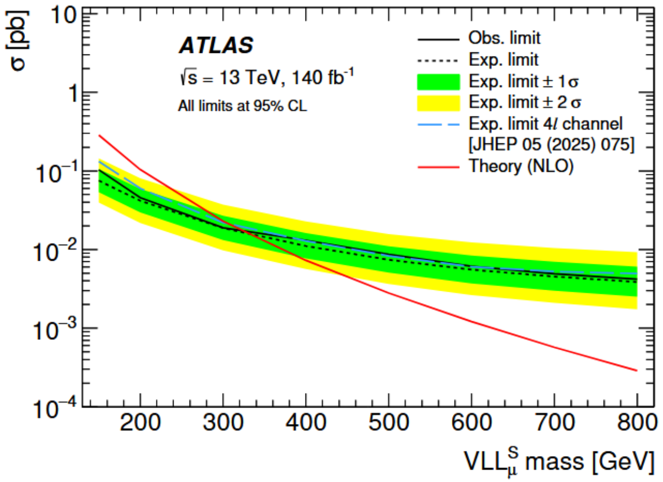
VLLμ (doublet)



VLLe (singlet)



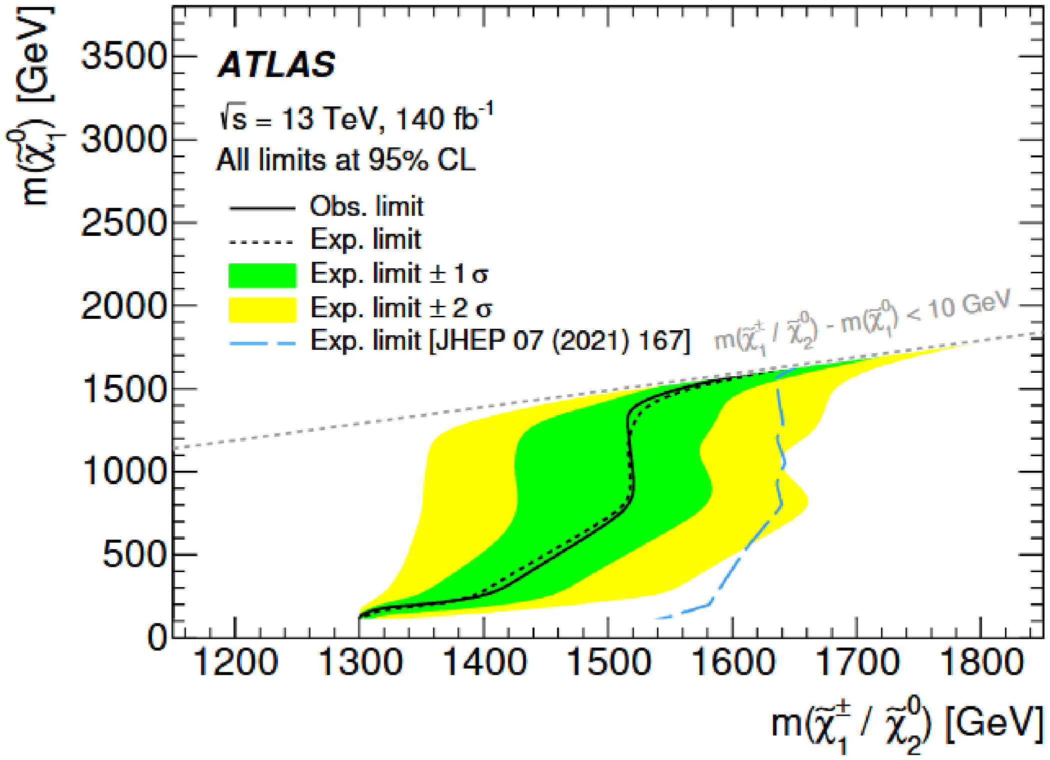
VLLμ (singlet)



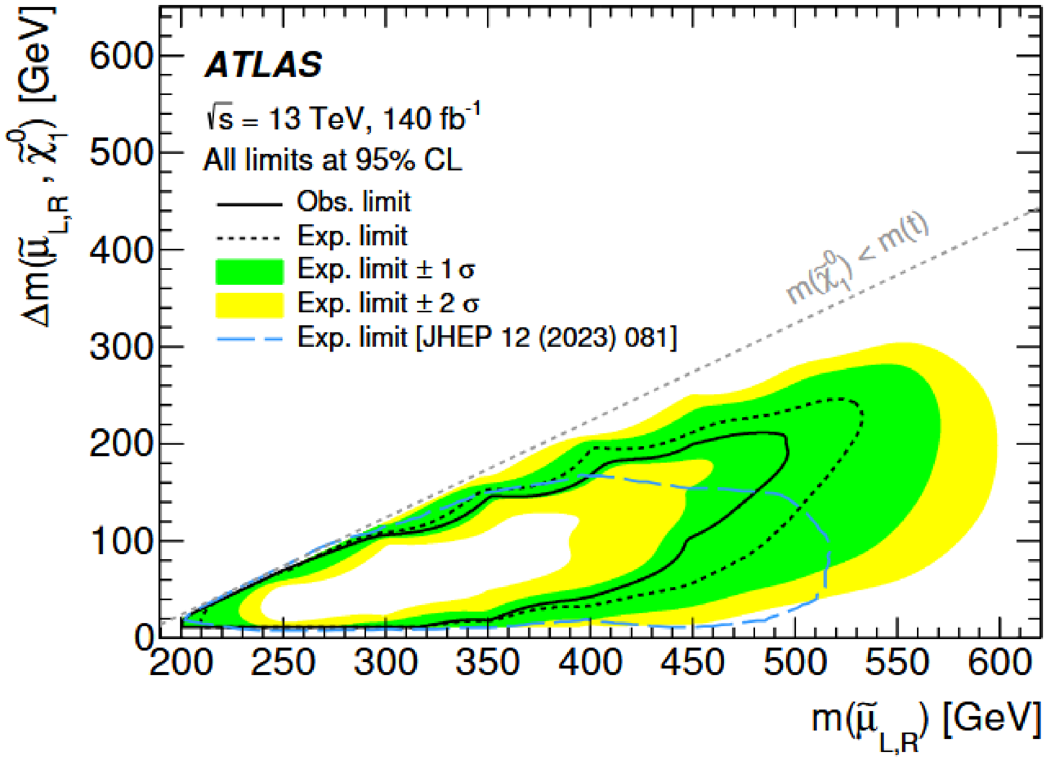
Model-dependent benchmark limits

- While performance is generally competitive, there are instances where it is expectedly weaker, or where the search is simultaneously less and more sensitive in different regions of the mass phase space.

Wino



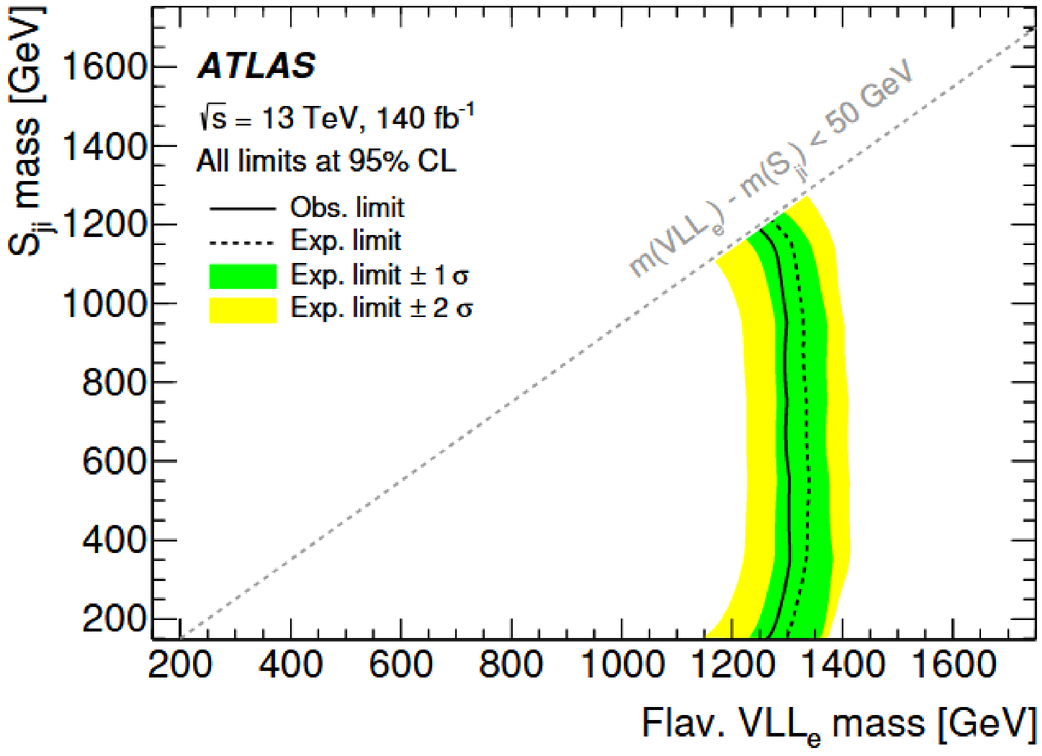
Smuon



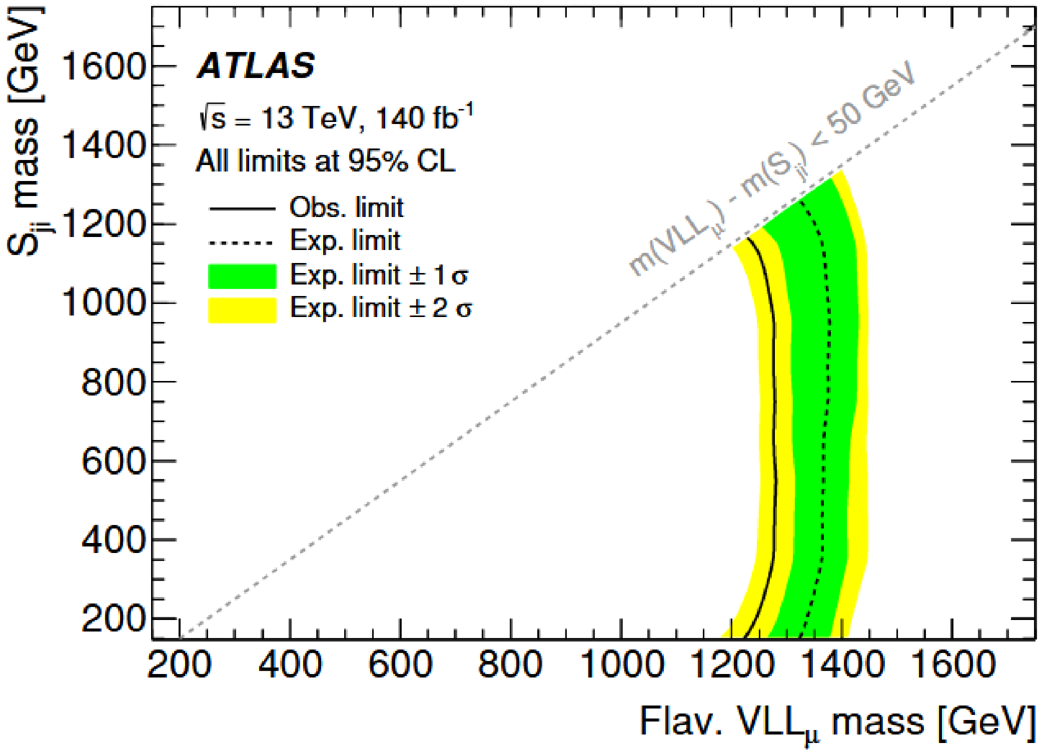
Model-dependent exclusion limits

- The **first LHC limits** are presented for the **Flavourful VLL model**, determined using 10000 S+B and B-only toys. The result exclude up to **1.3(1.25) TeV** for the Fl. VLLe(μ).

Flavourful VLLe



Flavourful VLLμ



Conclusions

- The full model-independent multilepton analysis is presented, marking the **first time anomaly detection** has been exploited in **multilepton** final states.

- **No significant excess** observed, with the largest local excess around 2σ found in the Q=0 0Z 2SFOS >99.9% and Q=±2 >90% regions.

- The analysis demonstrates **competitive sensitivity** across a range of signal benchmark models.

- An **LHC-first exclusion limit** was set on the Flavourful VLL model excluding up to 1.3 (1.25) TeV for Flavourful VLLe(μ).

- The analysis (**arXiv** here) represents a **stepping stone** for a much more ambitious ongoing search using **Run 3** data.

BACKUP

Object definition

- We use SLT in 2ℓSS and 3ℓ, DLT in $\geq 4\ell$, Standard Overlap Removal, and apply the conversion veto everywhere (DFCA < 1).

Feature	Criterion
Identification	`LooseAndBLayer likelihood`
Isolation	`FCLoose`
Pseudorapidity range	$(\|\eta\| < 1.37)$ || $(1.52 < \|\eta\| < 2.47)$
Energy calibration	`es2018_R21_v0` (ESModel)
Transverse momentum	$p_T > 10$ GeV
Object quality	Not from a bad calorimeter cluster (`BADCLUSELECTRON`) Remove clusters from regions with EMEC bad HV (2016 data only)
Track to vertex association	$\|d_0^{BL}(\sigma)\| < 5$ $\|\Delta z_0^{BL} \sin\theta\| < 0.5$ mm

Table 6: Electron selection criteria.

Feature	Criterion
Selection working point	`Loose`
Isolation	`FCLoose`
Veto	`BADMUONVETO`
Momentum calibration	Sagitta correction used
Transverse momentum	$p_T > 10$ GeV
$\|\eta\|$ cut	< 2.5
d_0 significance cut	3
z_0 cut	0.5 mm

Table 7: Muon selection criteria.

Parameter	Value
Algorithm	Calo-based
Soft term	Track-based (TST)
MET operating point	`Tight`
Selection requirements	
Observable	Requirement
E_T^{miss}	No cut
$\sum E_T / E_T^{miss}$	No cut
Object-based E_T^{miss} significance	No cut

Table 11: E_T^{miss} reconstruction criteria.

Feature	Criterion
Algorithm	Anti-k_T
R-parameter	0.4
Input constituent	PFlow
Analysis release number	21.2.147
`CalibArea` tag	
Calibration configuration	`JES_MC16Recommendation_Consolidated_EMTopo_Apr2019_Rel21.config`
Calibration sequence (Data)	`JetArea_Residual_EtaJES_GSC_Insitu`
Calibration sequence (MC)	`JetArea_residual_EtaJES_GSC_Smear`
Selection requirements	
Observable	Requirement
Jet cleaning	`LooseBad`
BadBatMan cleaning [52]	No
p_T	> 25 GeV
$\|\eta\|$	< 2.5
JVT	> 0.5 for $p_T < 60$ GeV, $\|\eta\| < 2.4$

Table 8: Jet reconstruction criteria.

Feature	Criterion
	PFlow Jets
Jet collection	`AntiKt4EMPFlowJets`
Jet selection	$p_T > 25$ GeV $\|\eta\| < 2.5$ same JVT cuts as for jets
Algorithm	`DL1r`
Operating point	Pseudo-continuous Eff = 77%
CDI	`2023-21-13TeV-MC16-CDI-2023-07-18_v1`

Table 9: b-tagging selection criteria.

Monte Carlo samples

Background samples

Process	DSIDs	Name
VV	364250-364255	Sherpa_222_NNPDF30NNLO_[llll-lvvv]
	364283-364284	Sherpa_222_NNPDF30NNLO_[llll-lvvv]_EW6
	364288-364290	Sherpa_222_NNPDF30NNLO_lll[lv]_lowMllPtComplement
	345705-345706	Sherpa_222_NNPDF30NNLO_ggllll_[0-130]M4l
	363355-363360	Sherpa_221_NNPDF30NNLO_[Z,W]qqZ[vv,ll]
VVV	364242-364249	Sherpa_222_NNPDF30NNLO_[WWW-ZZZ]_[]l[]v_EW6
VH	346310-346312	PowhegPythia8EvtGen_NNPDF30_AZNLO_*
Z+jets	700320-700337	Sh_2211_Z[ee,mm,tt,vv]
W+jets	700338-700349	Sh_2211_W[ev,mv,tv]
Z+jets IntC	346413	PhPy8EG_AZNLOCTEQ6L1_ZmumuWithInternalConversionFilter
Z+jets ExtC	361107	PowhegPythia8EvtGen_AZNLOCTEQ6L1_Zmumu
$t\bar{t}(Z/\gamma^*)$	504330,4,8	aMCPy8EG_NNPDF30NLO_A14N23LO_tt[ee,mm,qq]
	410276-410278	aMcAtNloPythia8EvtGen_MEN30NLO_A14N23LO_tt_mll_1_5
	(504329,33,41)	(aMCH7EG_NNPDF30NLO_H721UE_tt[ee,mm,tt])
$t\bar{t}W$	501720	aMCPy8EG_A14NNPDF23LO_ttW_FxFx01jNLO
	412123	MGPy8EG_A14_NNPDF23LO_EWttWsm
	(410155)	(aMcAtNloPythia8EvtGen_MEN30NLO_A14N23LO_ttW)
$t\bar{t}H$	346343-346345	PhPy8EG_A14NNPDF23_NNPDF30ME_ttH125_*
	(346443-346445)	(aMcAtNloPythia8EvtGen_ttH_noShWe_*)
$4t$	412043	aMcAtNloPythia8EvtGen_A14NNPDF31_SM4topsNLO
	(700046)	(Sh_2210_tttt)
$3t$	304014	MadGraphPythia8EvtGen_A14NNPDF23_3top_SM
$t\bar{t}WW$	410081	MadGraphPythia8EvtGen_A14NNPDF23_ttbarWW
$t\bar{t}VV$	500460-500462	MGPy8EG_A14NNPDF23LO_tt[ZZ,WH,HH]
$t\bar{t}$	410470	PhPy8EG_A14_ttbar_hdamp258p75_nonallhad
	410397-410399	ttbar_wb[ee,mm,tt]_MEN30LO_A14N23LO
tWZ	410408	aMcAtNloPythia8EvtGen_tWZ_Ztoll_minDR1
tZ	410560	MadGraphPythia8EvtGen_A14_tZ_4fl_tchan_noAllHad
tW	410646-410647	PowhegPythia8EvtGen_A14_Wt_DR_inclusive_*
t	410644-410645	PowhegPythia8EvtGen_A14_singletop_schan_lept
	410658-410659	PhPy8EG_A14_tchan_BW50_lept_[anti]top

Signal samples

Process	DSIDs	Name
VLLe Doublet	512071-512083	MGPy8EG_NNPDF31_MVLL[150-1300]_el_doublet
VLLe Singlet	512084-512096	MGPy8EG_NNPDF31_MVLL[150-1300]_el_singlet
VLLμ Doublet	512097-512109	MGPy8EG_NNPDF31_MVLL[150-1300]_mu_doublet
VLLμ Singlet	512110-512122	MGPy8EG_NNPDF31_MVLL[150-1300]_mu_singlet
VLLe with S	518405-518425	MGPy8EG_VLL_el_MVLL[200-1200]_2Scal_MScal[150-1150]
VLLμ with S	518426-518446	MGPy8EG_VLL_mu_MVLL[200-1200]_2Scal_MScal[150-1150]
$VH \to hh$	517565-517572	MGPy8EG_AZallHhh_mH[300-1000]
Wino	397503-397526	MGPy8EG_A14N23LO_Wino_[1300-1900]_[10-1890]_LLE12k
	397527-397546	MGPy8EG_A14N23LO_Wino_[800-1300]_[10-1290]_LLEi33
slepton	397813-397830	MGPy8EG_A14N23LO_LV_[900-1300]_[10-1290]_LLE12k
	397527-397546	MGPy8EG_A14N23LO_LV_[700-1000]_[10-990]_LLEi33
gluino	449876-449882	MGPy8EG_A14N23LO_GG_[2000-2200]_[100-1990]_LLE12k
	449883-449895	MGPy8EG_A14N23LO_GG_[1600-2200]_[100-2190]_LLEi33
smuon	508523-508546	MGPy8EG_A14N23LO_SmuonSmuon_directRPVLQD_[200-400]_[180-390]
	512134-512157	MGPy8EG_A14N23LO_SmuonSmuon_directRPVLQD_[250-550]_[200-530]

Anomaly detection technique

- Method requires:
  - Same input and output dimensionality.
  - Invertible learnt mapping.
  - Tractable determinant of the Jacobian.

- Focus on the Real Non-Volume Preserving (RealNVP) model, separating z into two disjoint subsets, z_1 and z_2 and then apply two neural networks (s_θ , m_θ):

$$x_1 = z_1,$$

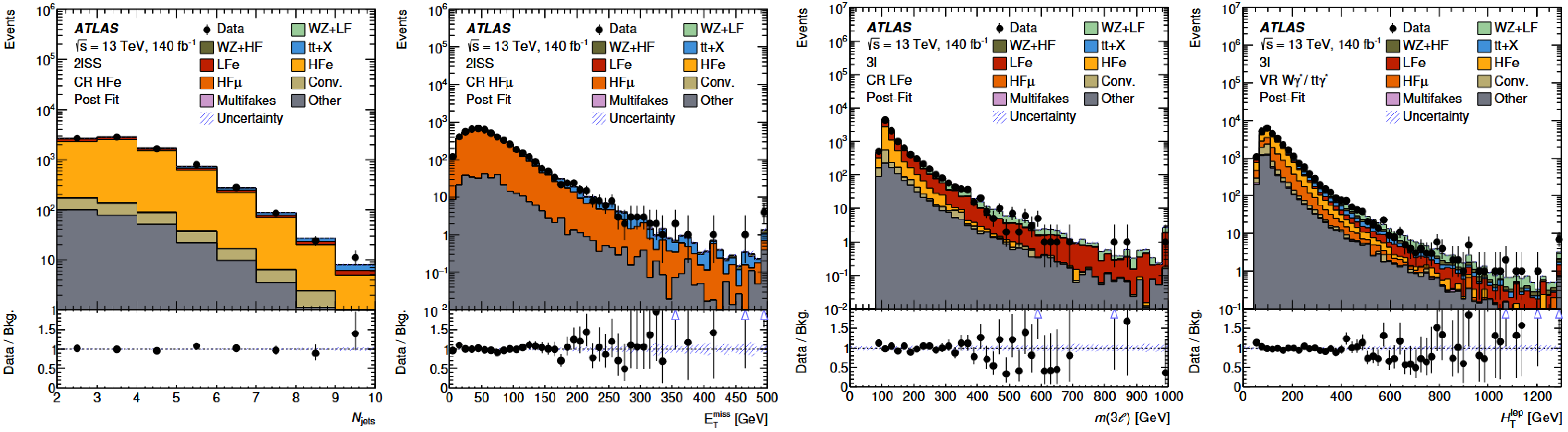
$$x_2 = e^{s_\theta(z_1)} + m_\theta(z_1),$$

- We then scale the probability to build an anomaly score in the [0, 1] range, according to:

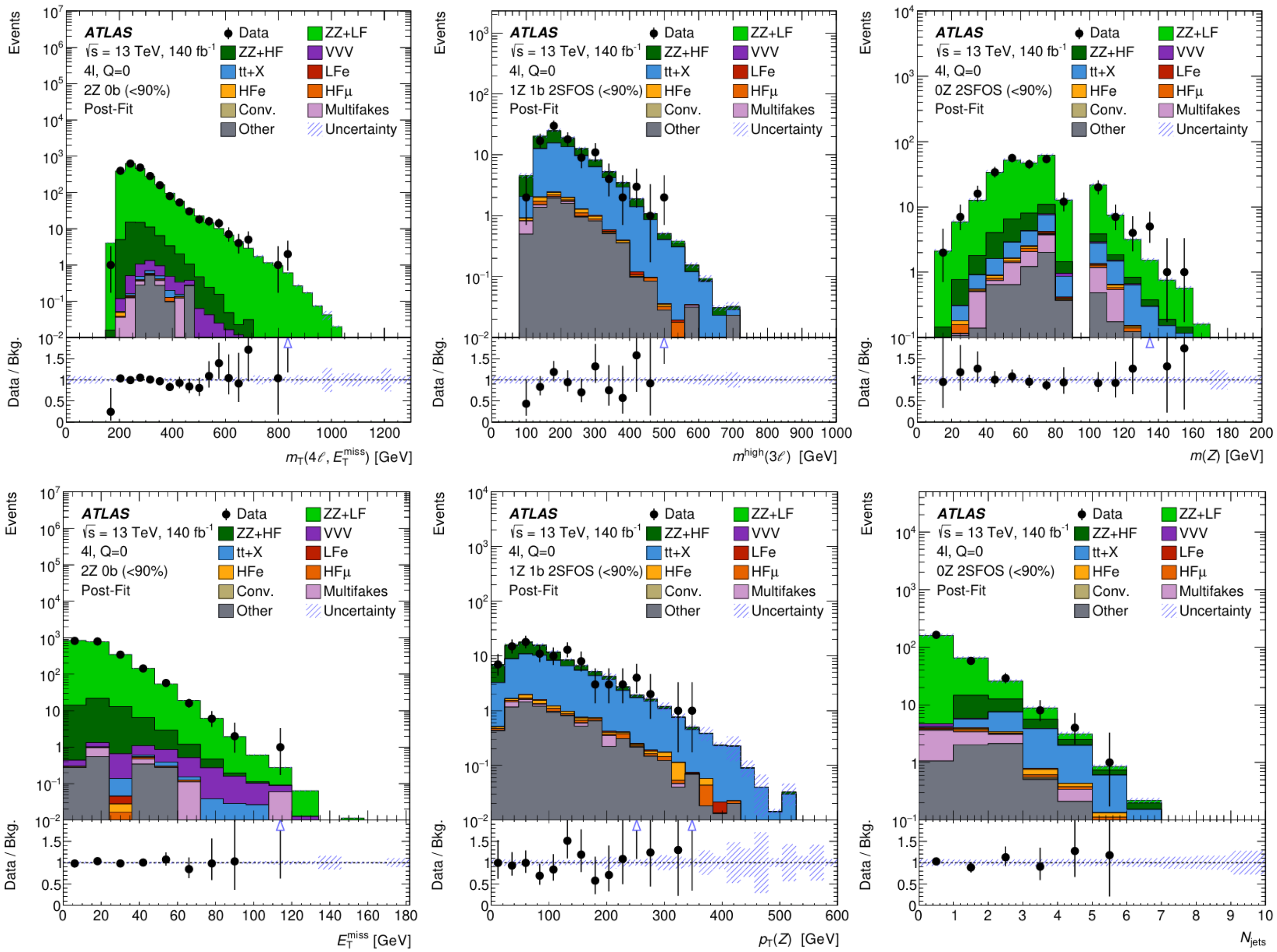
$$s(x) = \frac{\log p(x) - \log p_{max}}{\log p_{min} - \log p_{max}}$$

- Anomaly score is then based only on background probability.

Fake modelling



Input variable data/MC agreement



Data-driven QmisID estimation

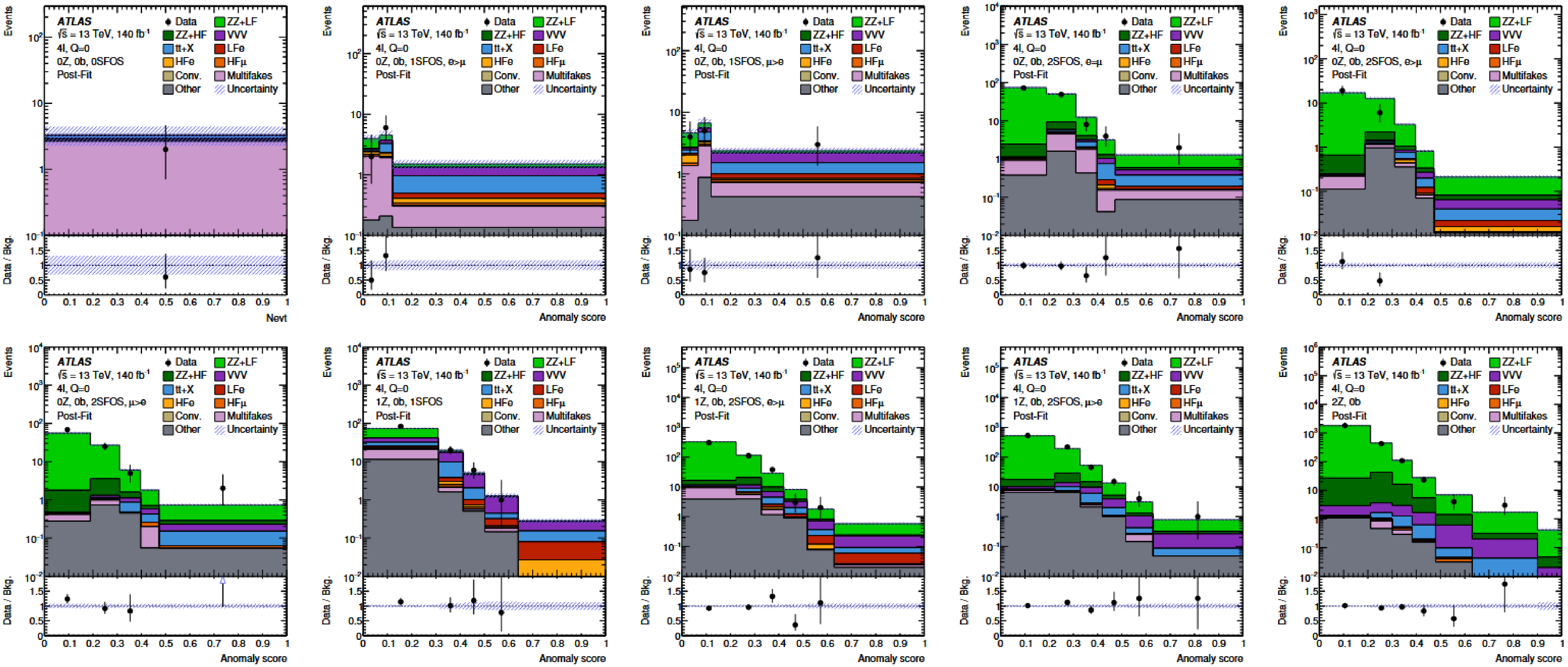
- Significant charge-misID contribution in 4ℓ Q=$\pm$2, which we estimate using a data-driven estimate derived from $Z\rightarrow ee$ events in $2\ell SS$. We use a likelihood technique to measure a QmisID rate in bins of p_T and η, and apply it to 4ℓ Q=0 data to derive a data-driven estimate.

- We then derive uncertainty measurements on the measured rates through 3 sources:

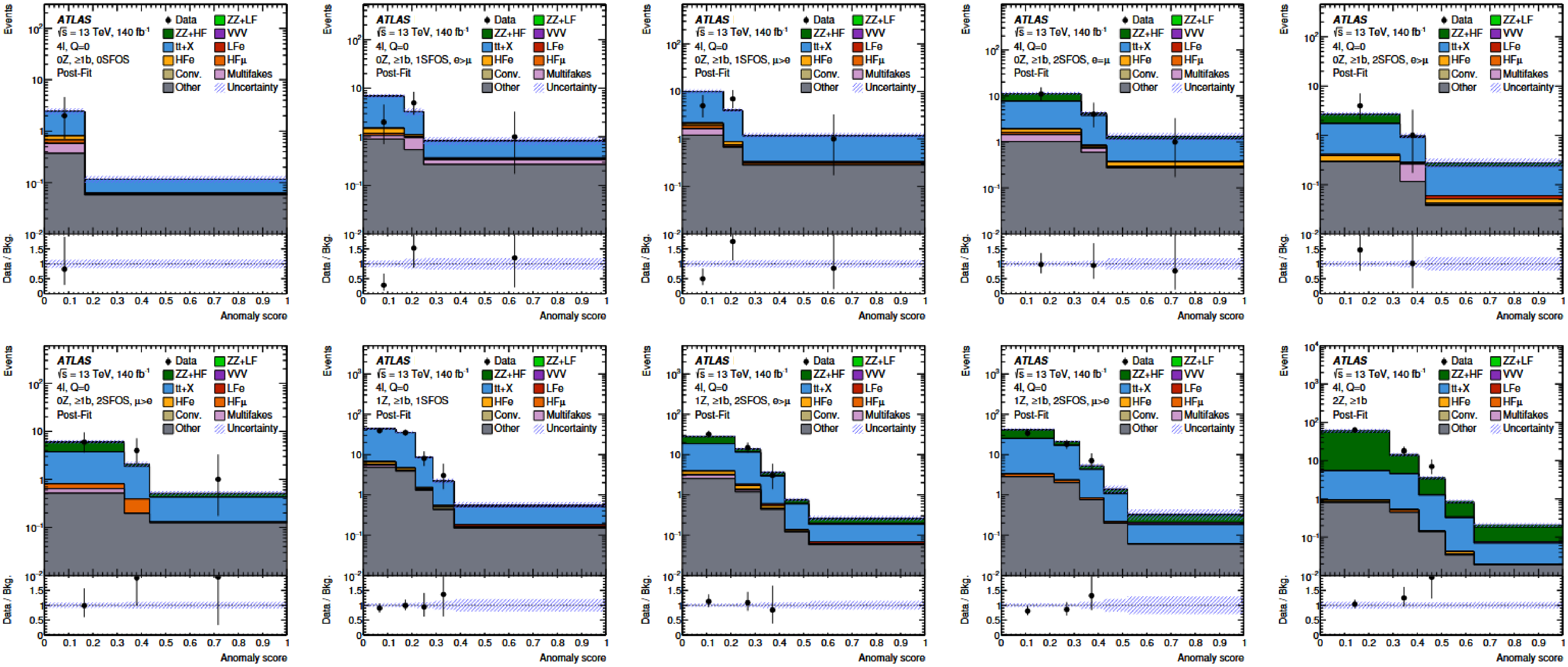
  - Error estimate from the likelihood.
  - Error estimate through difference in rates with truth matched $Z\rightarrow ee$ events.
  - Variation of rate within M_z window.

- We produce a final uncertainty by adding the three sources in quadrature, and add it to the systematics model.

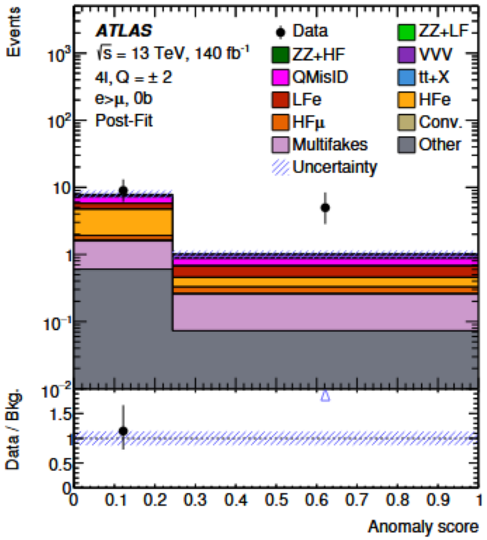
Model-dependent signal regions, 4ℓ Q=0 0b

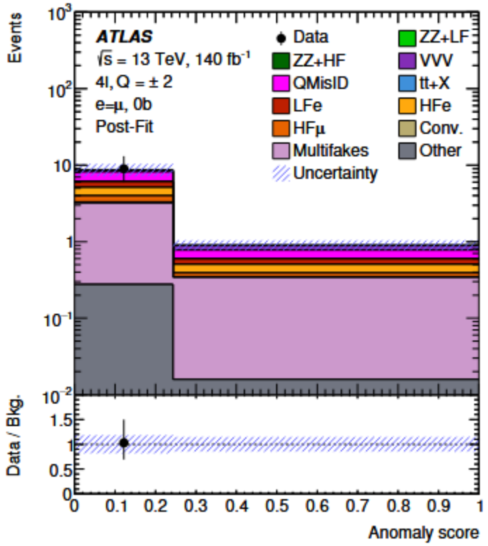


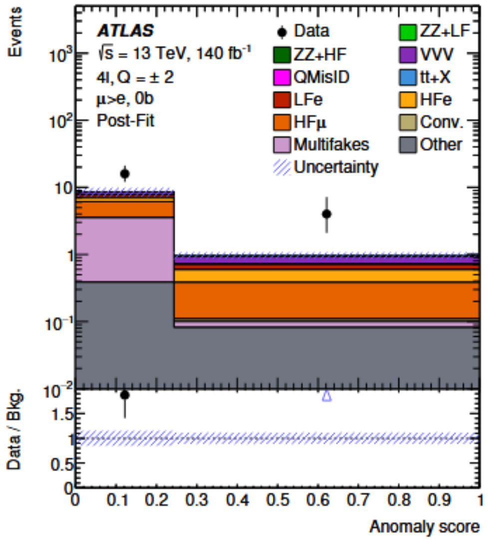
Model-dependent signal regions, 4ℓ Q=0 ≥1b

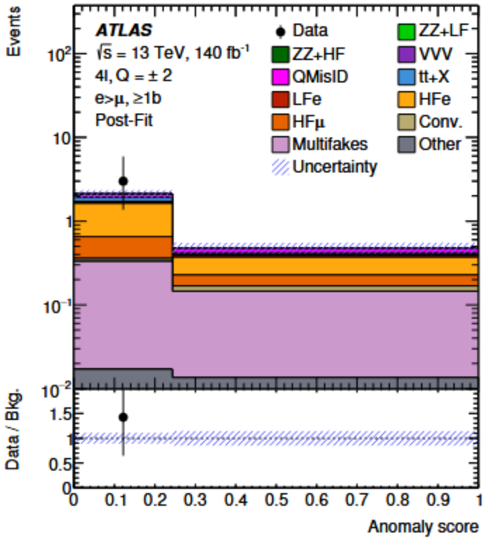


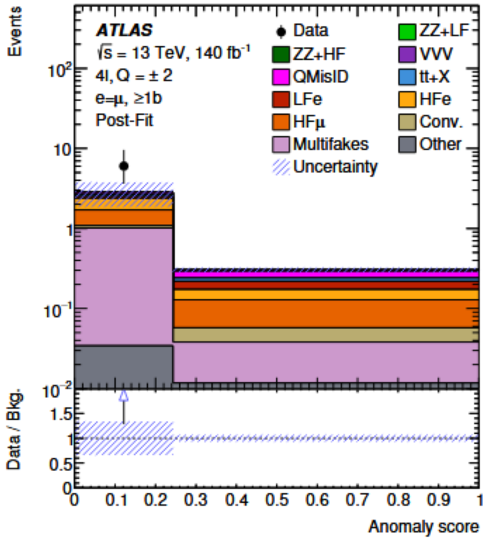
Model-dependent signal regions, 4ℓ Q=±2

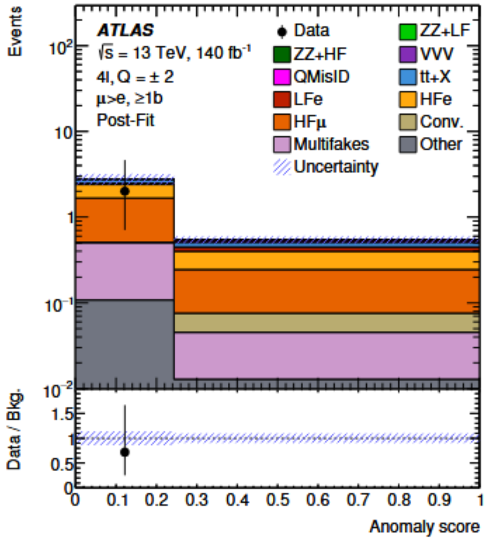












Model-dependent signal regions, ≥ 5ℓ