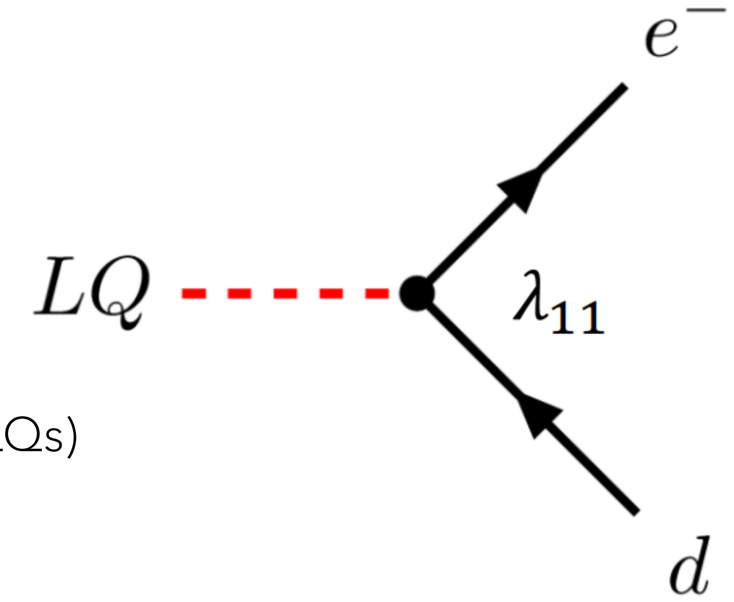
Search for third generation LQs using parametrized neural networks

VIKTOR VASILEV

Theoretical Background

- LQs are predicted by some Grand Unified Theories that aim to unify the strong with the electroweak force at high energies
- Carry both color and charge
- Have non-zero baryon and lepton number
- Decay into lepton-quark pair
- Their charge can be 1/3, 2/3, 4/3, 5/3
- Their spin can be either 0 (scalar LQs) or 1 (vector LQs)

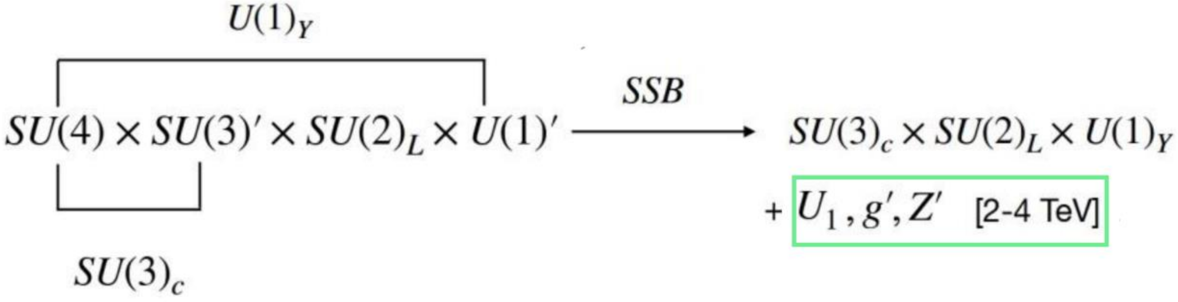


Scalar LQs

- They can be added ‘easily’ to the current Standard Model
- No need for introduction of larger theoretical framework, addition of Yukawa terms to the Lagrangian of the SM

Vector LQs

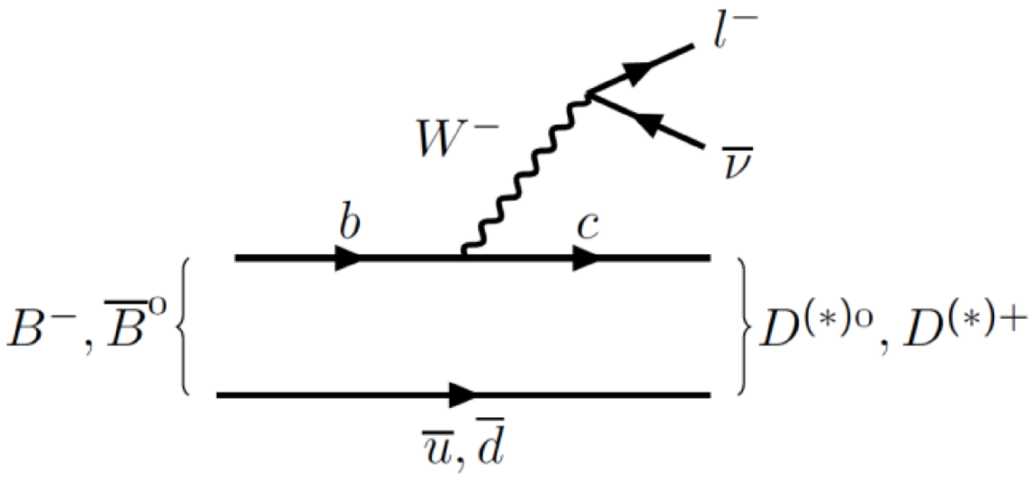
- Introduce another gauge interaction
- Need bigger symmetry group to explain them
- Two main different scenarios, depending on the strength of the coupling -> “Minimal Coupling” and “Yang-Mills”
- The theories predict the existence of more heavy-mass particles, e.g. couloron g’; Z’

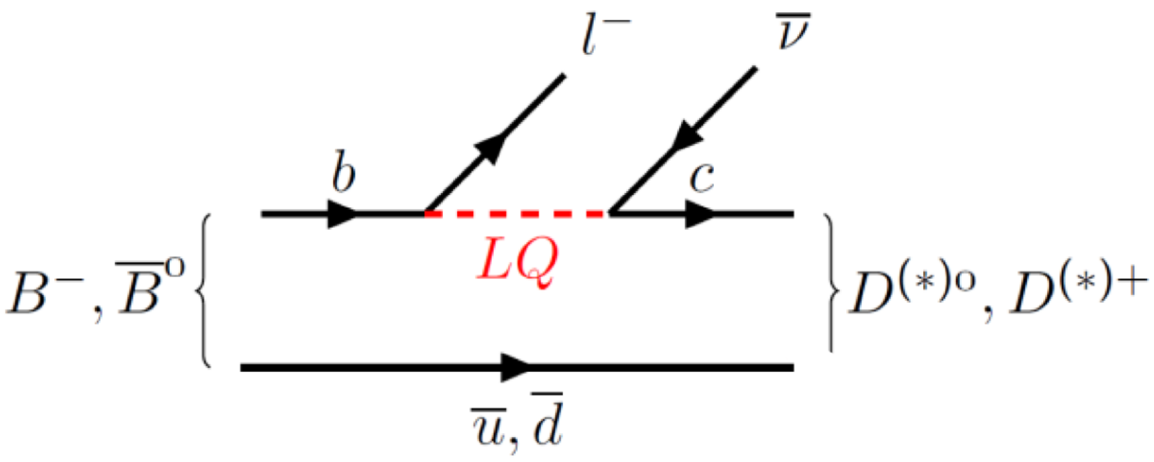


Lepton Flavor Violation

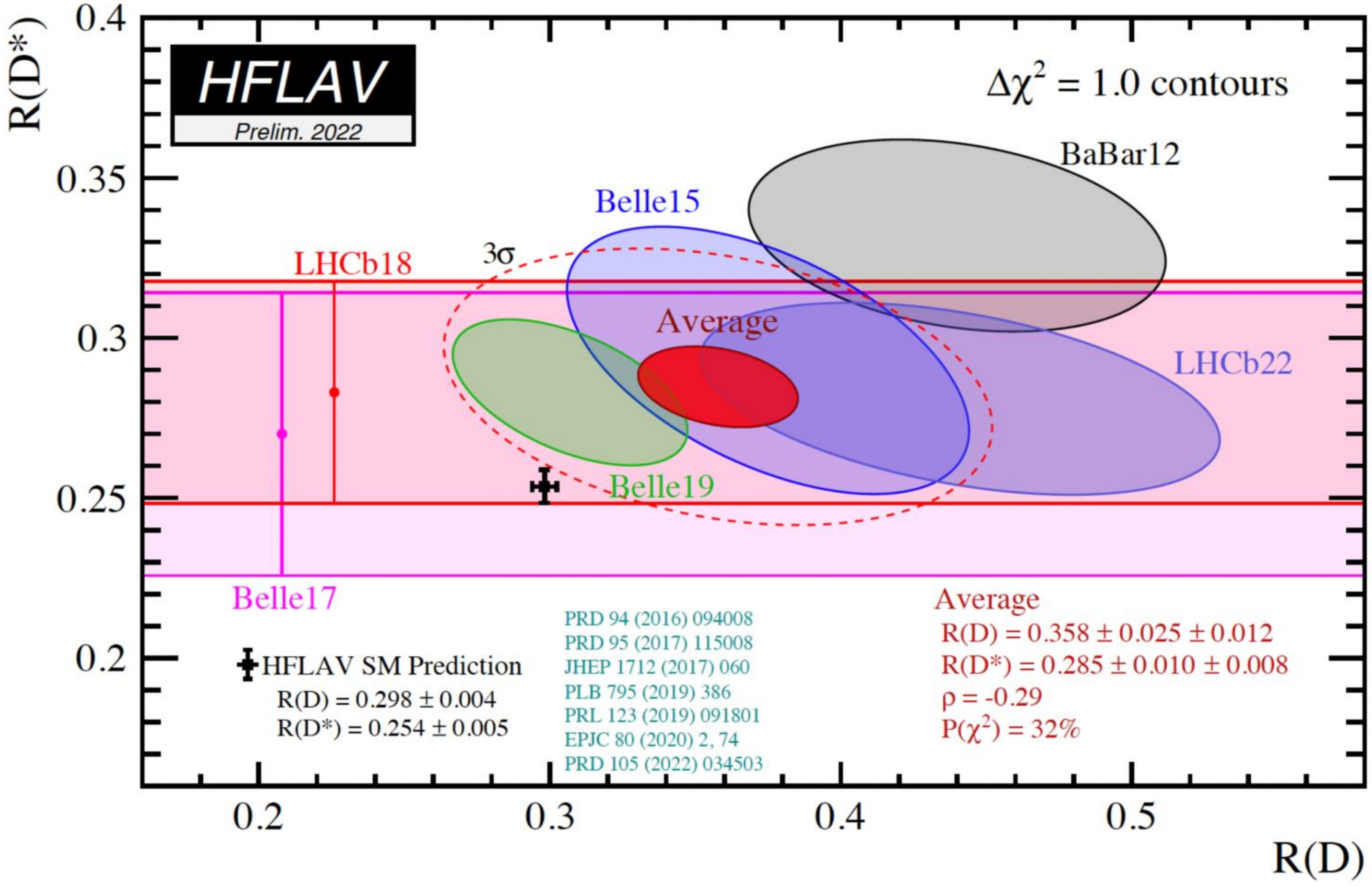
- Multiple experiments have shown deviations from the SM, current global average is 3.2σ
- The quantity that deviates from the prediction is:

$$R = \frac{\mathrm{B}(B \rightarrow D^{(*)}\tau\nu_{\tau})}{\mathrm{B}(B \rightarrow D^{(*)}l\nu_{\tau})}$$





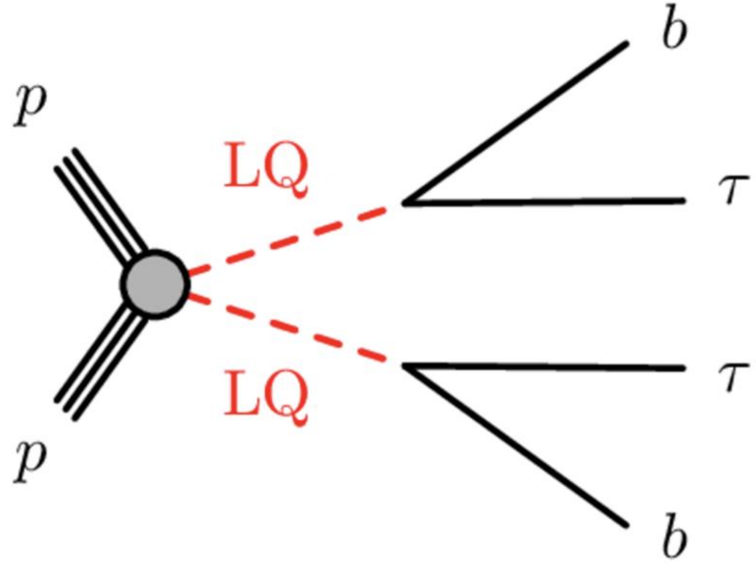
Lepton Flavor Violation



Comparison of results for $R(D^*)$ and $R(D)$ of different experiments with the prediction of the SM

What does the paper focus on?

- Focus on third generation up-type LQ pair, which decays into $\mathrm{b}\tau$
- Before this paper, up-type LQs (decaying into $b\tau$ or $t\nu$) were excluded for masses below $1150\ GeV$ for all values of the branching ratio
- Hypothetically LQs coupling harder to τ-leptons could explain the flavor anomaly
- The mass range considered for the LQs is 300 GeV to 2000 GeV



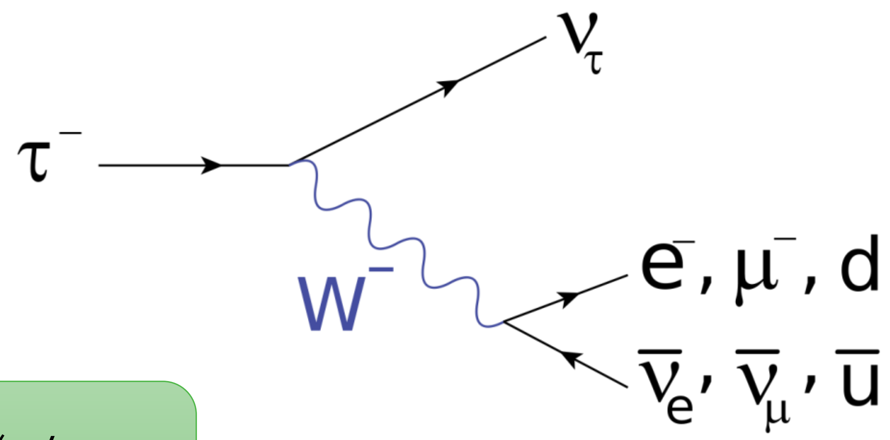
What does the paper look for?

- The analysis signature is two jets, at least one of which must be identified as containing a b-hadron; and two τ-leptons
- We define two different channels $\tau_{\text{lep}}\tau_{\text{had}}$ ($\sim$46%) and $\tau_{\text{had}}\tau_{\text{had}}$ ($\sim$42%), depending on whether one or zero of the τ-leptons decay leptonically

τ-decays

64.8% → Hadronic decay (Combination of pions + neutrinos)

35.2% → Leptonic decay (e/μ + neutrinos)



Event selection

- Muons and electrons each have different working points ('veto' -> 'loose'; 'signal'->'tight')
- Trade-off quality<-> number of candidates
- All light lepton candidates must also satisfy an isolation criterion – no other tracks in a fixed cone around their trajectory, resulting in $> 99\%$ efficiency in the Signal Region (SR)
- Boosted Decision Trees (BDT) and Recurrent Neural Networks (RNN) are used to distinguish the hadronically decaying τ-leptons

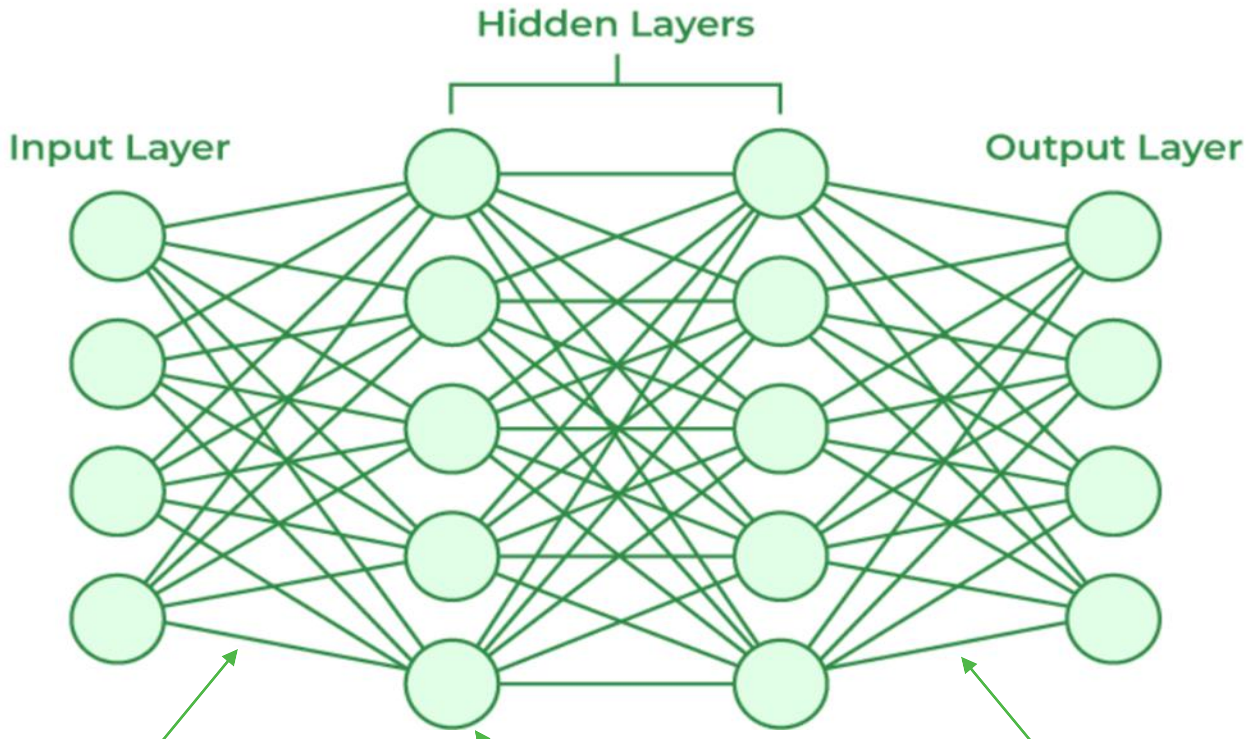
Signal Region

Selection \ Channel	$\tau_{\text{lep}}\tau_{\text{had}}$ channel	$\tau_{\text{had}}\tau_{\text{had}}$ channel
e/μ	Exactly 1 'signal/tight' e or μ $p_{\text{T}}^{e} > 25, 27$ GeV $p_{\text{T}}^{\mu} > 25, 27$ GeV	No 'loose' or 'tight' e or μ
$\tau_{\text{had-vis}}$	Exactly 1, oppositely charged to the e or μ $p_{\text{T}}^{\tau} > 100$ GeV	Exactly 2, with opposite charges $p_{\text{T}}^{\tau} > 100, 140, 180\ (20)$ GeV
Jets	≥ 2 jets, 1 or 2 b-jets, $p_{\text{T}}^{\text{jet}} > 25, 27$ GeV	
Additional requirements	$m_{\tau\tau}^{\text{MMC}}(\text{'rough' estimate}) \notin [40 - 150]$ GeV $E_{\text{T}}^{\text{miss}} > 100$ GeV $s_{\text{T}} = \sum_{2\tau+2jets+miss} \lvert p_T \rvert > 600$ GeV	

Selection criteria for the signal region

Short Introduction to Neural Networks

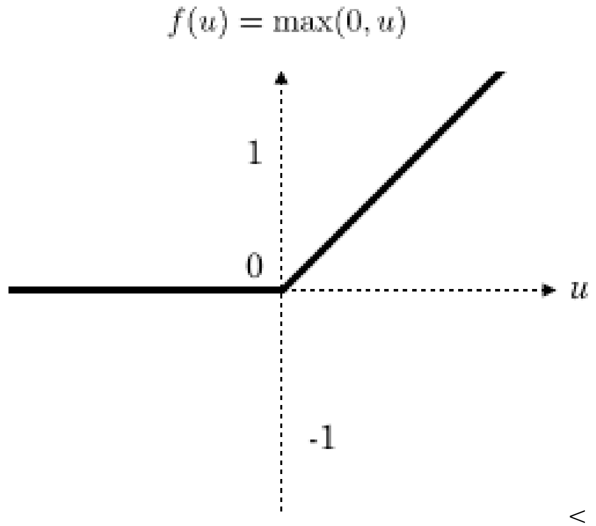
- Trying to fit a high-dimensional non-analytic 'function'



Weights on each edge

Non-linear Activation function

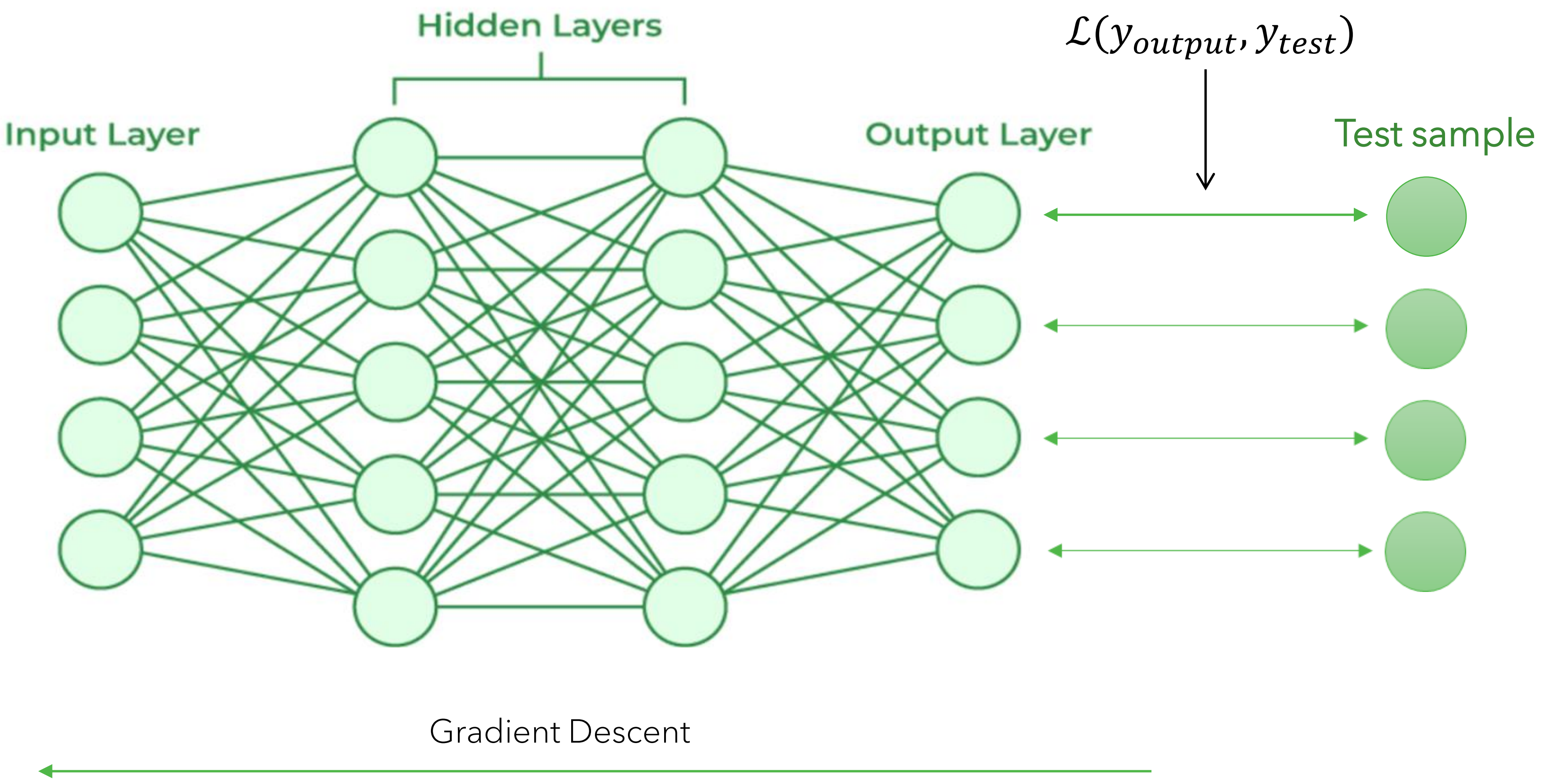
Matrix multiplication



RELU activation function

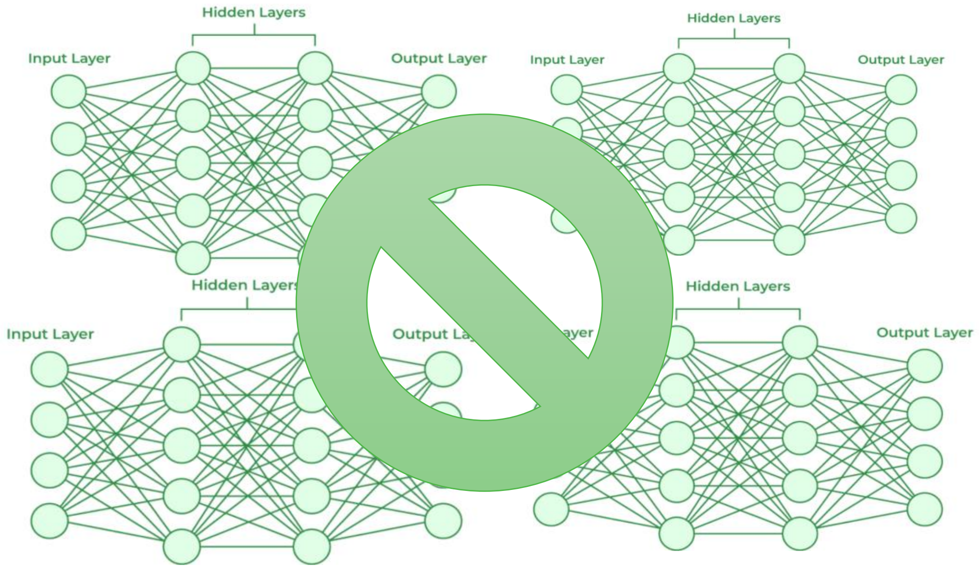
Short Introduction to Neural Networks

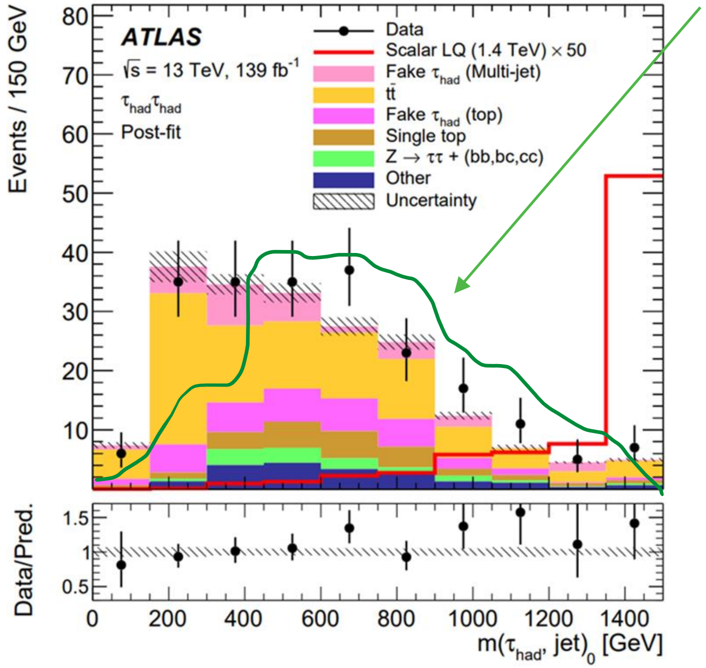
- Trying to fit a high-dimensional non-analytic 'function'



How do we tackle the problem with the unknown mass?

- "Brute Forcing" by training a lot of NNs on separate values of the LQ mass and interpolating for the values in-between
- Problems with that:
  - Not **smooth/continuous** range for the results, i.e. we have trained a lot of separate NNs on discrete values
  - Computationally expensive – each NN takes time and energy to train

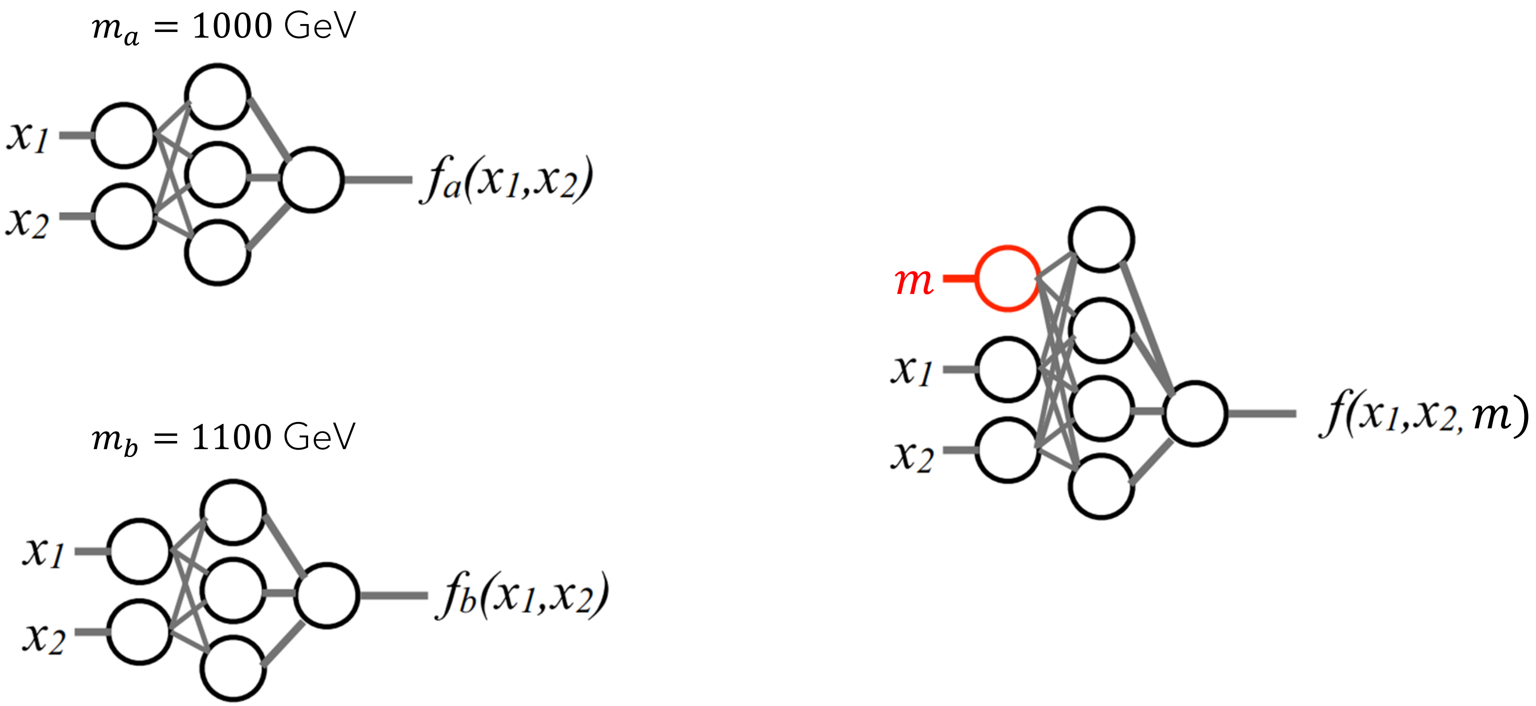




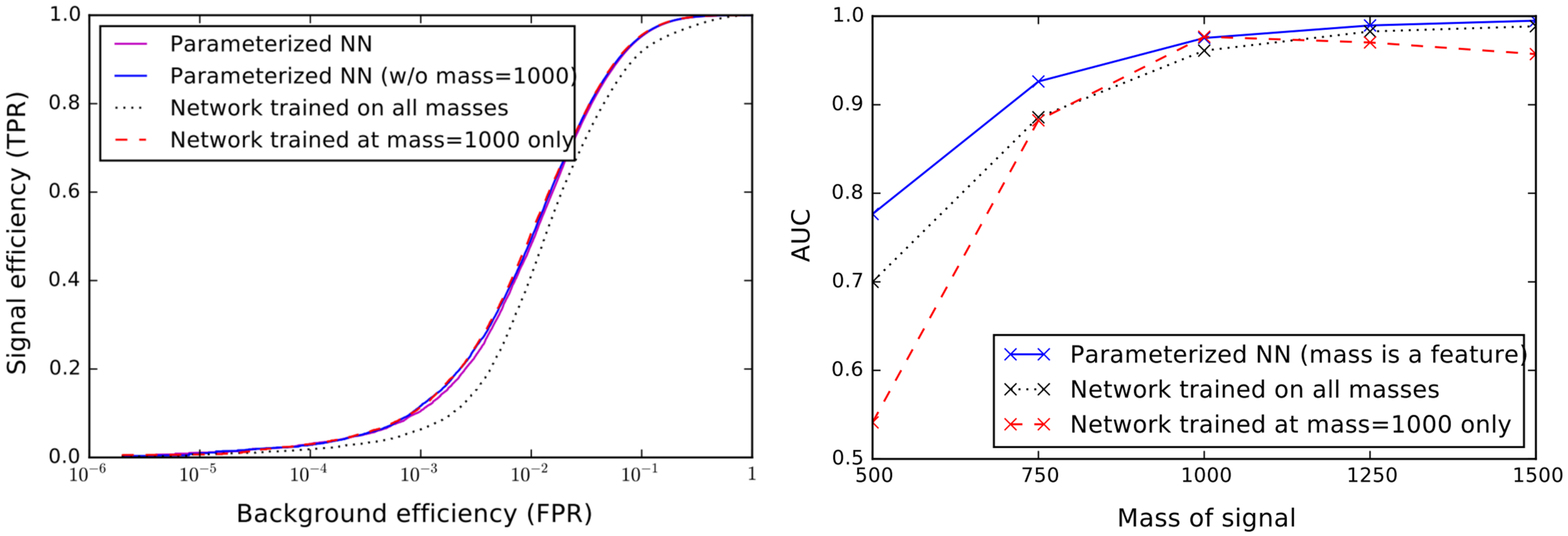
Signal for LQ with different mass

Parametrized Neural Network

- We have some distribution of the unknown parameter (in our case the mass of the LQ) as an input



Simulated Test of PNN



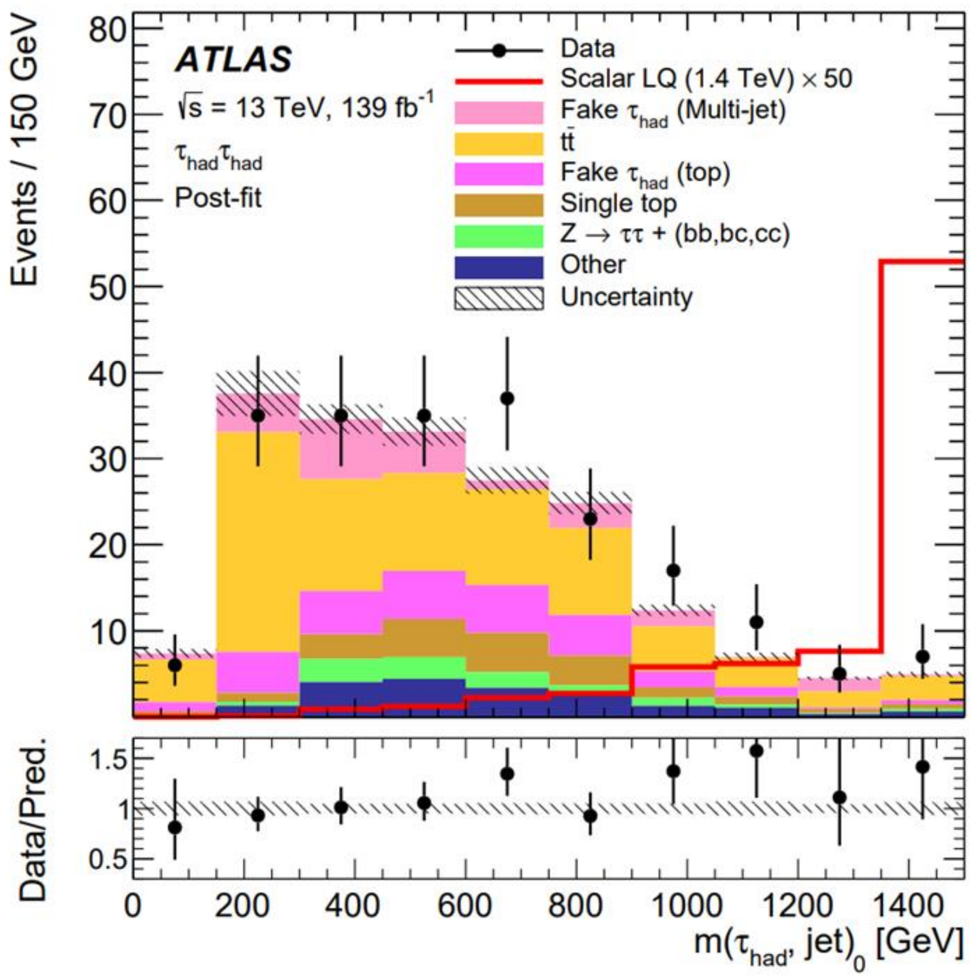
True positive vs false positive graph on the left and AUC score on the right. All are tested against a sample trained on $m_X = 1000$ GeV.

NNs are trained on $m_X = 500, 750, 1000, 1250, 1500$ GeV (except said otherwise)

(TPR = TP/(TP+FN) ; FPR = FP/(TN+FP)); AUC = Area Under Curve

Input parameters

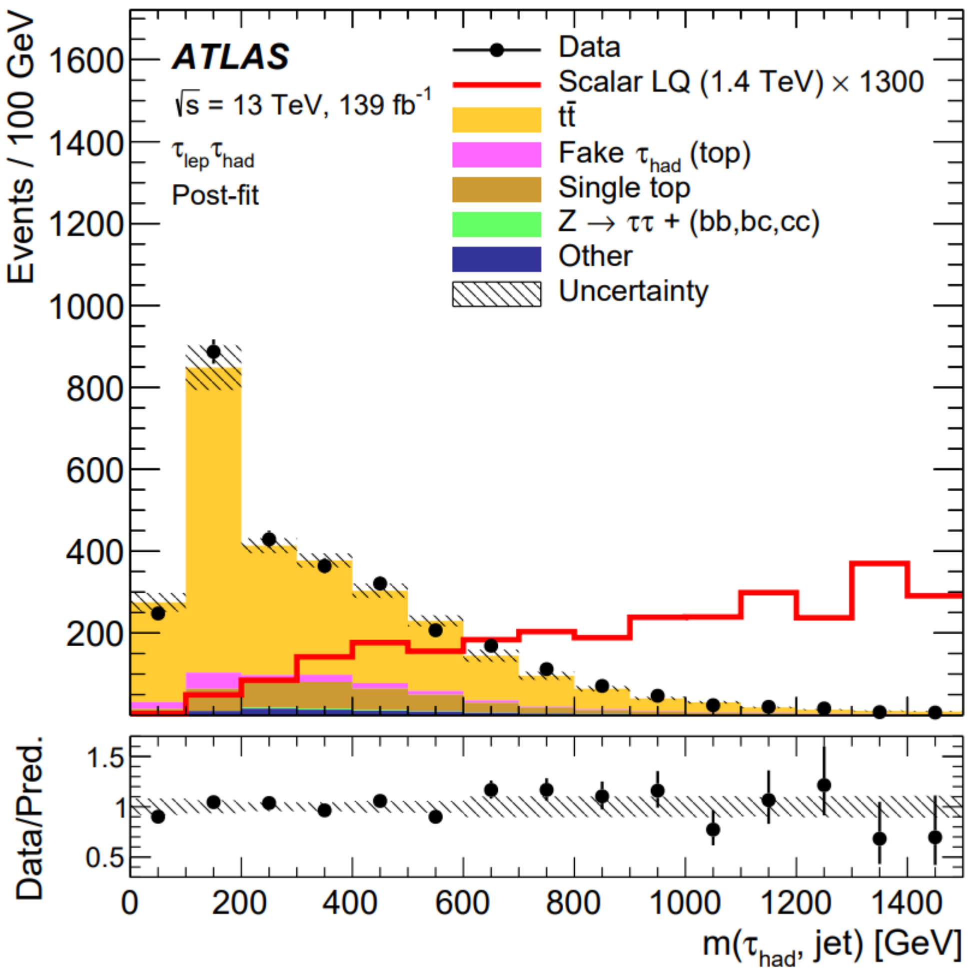
- $m(\tau, \text{jet})_{0,1}$ -> the mass of the larger (0) and the smaller (1) LQ candidate obtained via mass-pairing strategy that consists in minimizing the difference between the two possible candidates (full hadronic channel only)



Input distribution for $m(\tau, \text{jet})_0$ for the full hadronic channel; expected signal for scalar LQ at 1.4 TeV is also shown

Input parameters

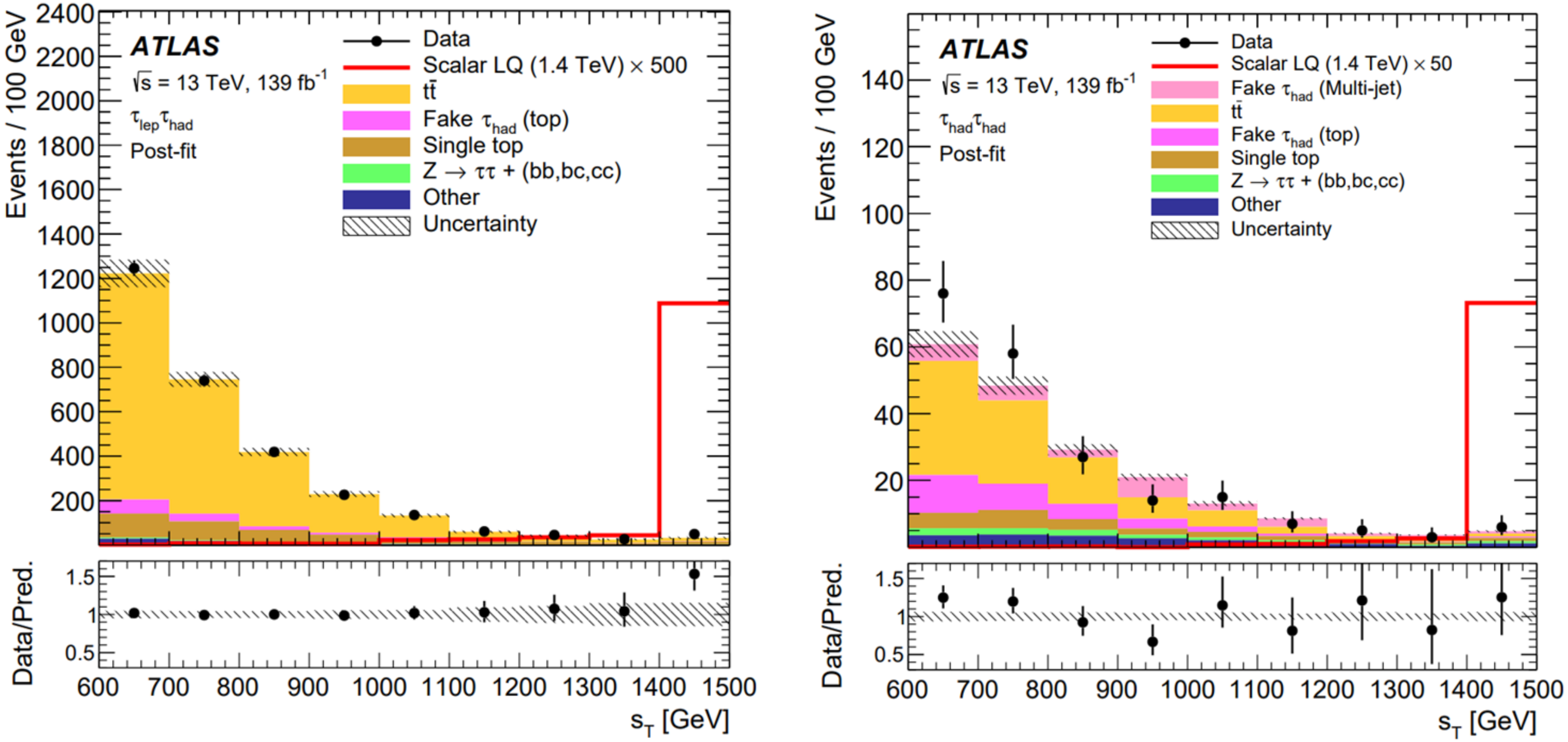
- $m(\ell,\ \text{jet}); m(\tau_{had}, \text{jet})$ -> mass of light lepton/$\tau_{had-vis}$ combined with the mass paired jet; only for the semileptonic channel



Input distribution for $m(\tau, \text{jet})$for the semileptonic channel; expected signal for scalar LQ at 1.4 TeV is also shown

Input parameters

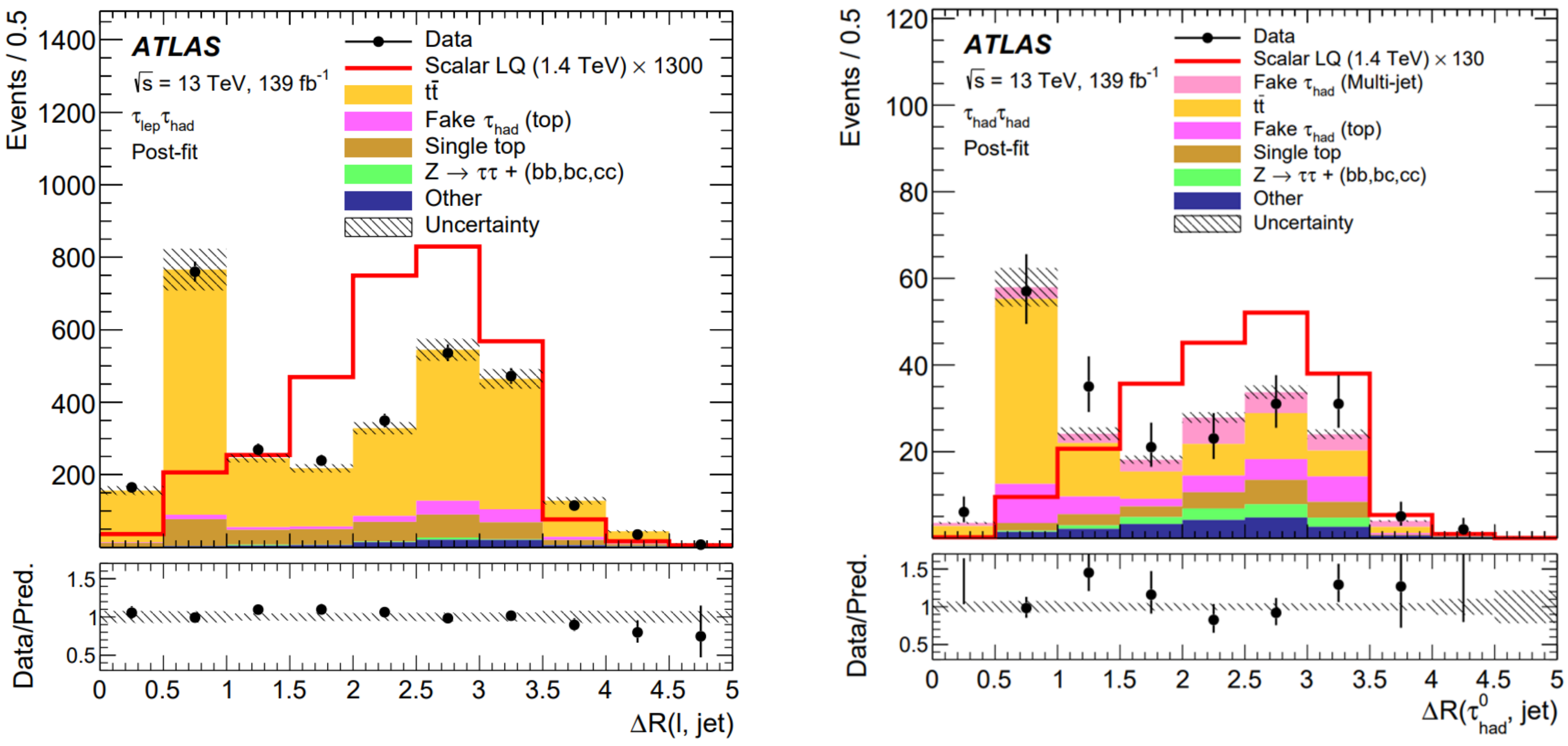
- s_T -> scalar sum of the transverse momenta of the two $\tau_{had-vis}$ (or just one + light lepton depending on the channel), the two leading jets and the missing transverse momentum



Input distribution in bins of s_T for the semileptonic channel (left) and full hadronic channel (right); expected signal for LQ at mass 1.4 TeV is also shown

Input parameters

$\Delta R(\ell, \text{jet})/\Delta R(\tau_{had}, \text{jet})$ -> Angular distance between the $\ell/\tau_{\text{had-vis}}$ (leading one) and the mass-paired jet



Input distribution in bins of angular distance for the semileptonic channel (left) and full hadronic channel (right); expected signal for scalar LQ at 1.4 TeV is also shown

Input features to the PNN

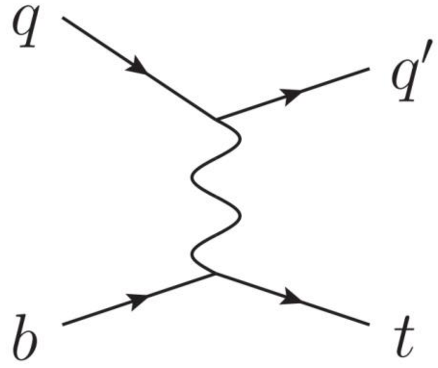
- $\tau_{\text{had-vis}}$ p_{T}^{0} -> transverse momentum of the highest-p_{T} $\tau_{\text{had-vis}}$
- N_{b-jets} -> number of b-jets
- $\Delta\phi\left(\ell, p_T^{miss}\right)$ -> azimuthal opening angle between the lepton and the missing transverse momentum (only for semileptonic channel)
- $p_T^{miss}\phi$ centrality -> quantifies how 'central' the missing momentum is with respect to the two τ particles / τ and light lepton

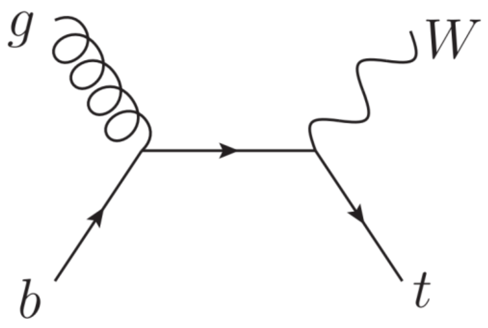
Input features to the PNN

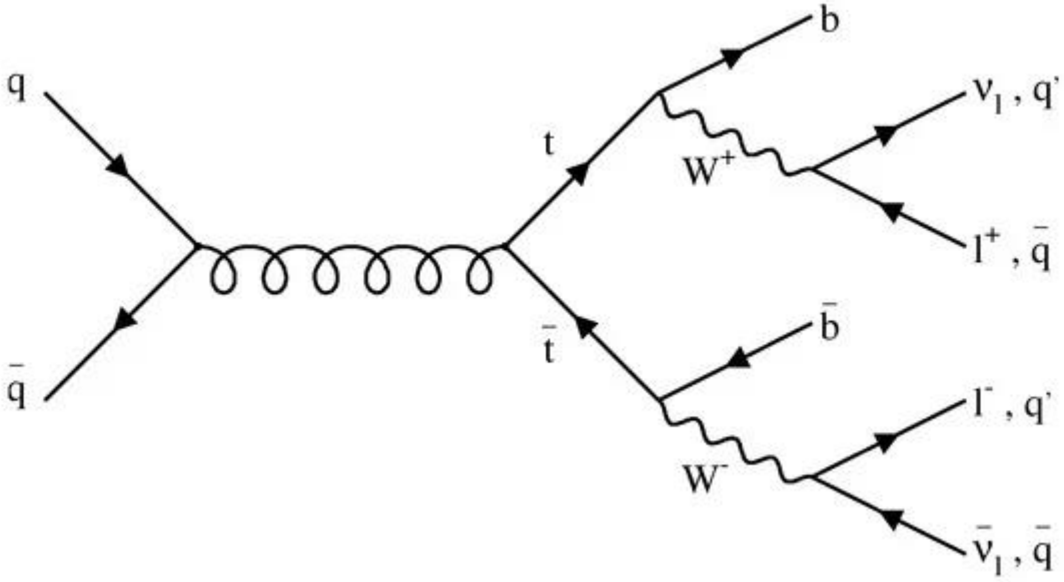
Variable	$\tau_{\text{lep}}\tau_{\text{had}}$ channel	$\tau_{\text{had}}\tau_{\text{had}}$ channel
$\tau_{\text{had-vis}}$ p_{T}^{0}	✓	✓
s_{T}	✓	✓
$N_{b-\text{jets}}$	✓	✓
$m(\tau, \text{jet})_{0,1}$		✓
$m(\ell, \text{jet})$, $m(\tau_{\text{had}}, \text{jet})$	✓	
$\Delta R(\tau, \text{jet})$	✓	✓
$\Delta\phi(\ell, E_{\text{T}}^{\text{miss}})$	✓	
$E_{\text{T}}^{\text{miss}}$ ϕ centrality	✓	✓

Background modelling

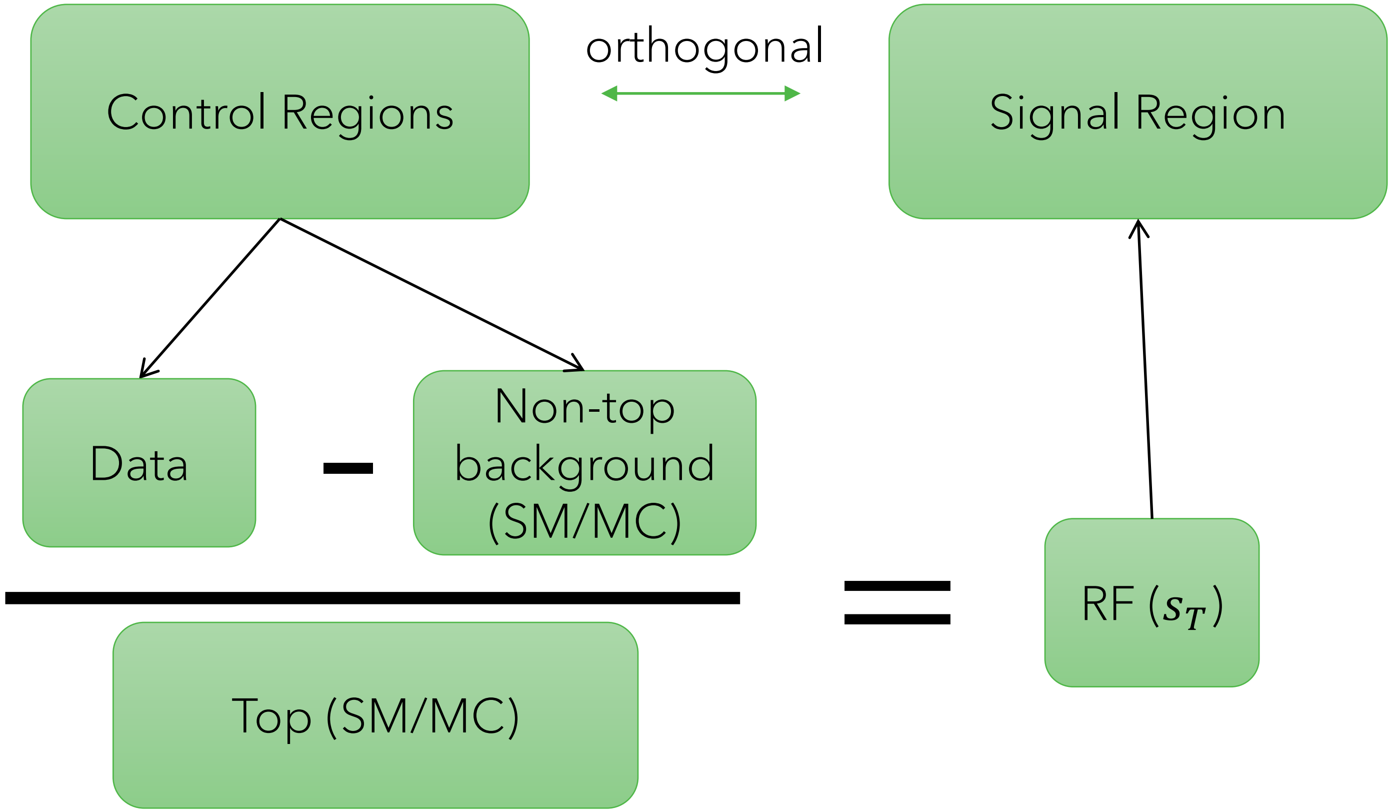
- Another very important part of this data analysis
- Dominant background is top production – both pair and single top-quark
- Subdominant background is Z boson production in association with heavy-flavor quarks (bb, bc, cc) (Z+HF)
- Multi-jet events in the full hadronic channel are also non-negligible
- Other backgrounds are estimated using simulation





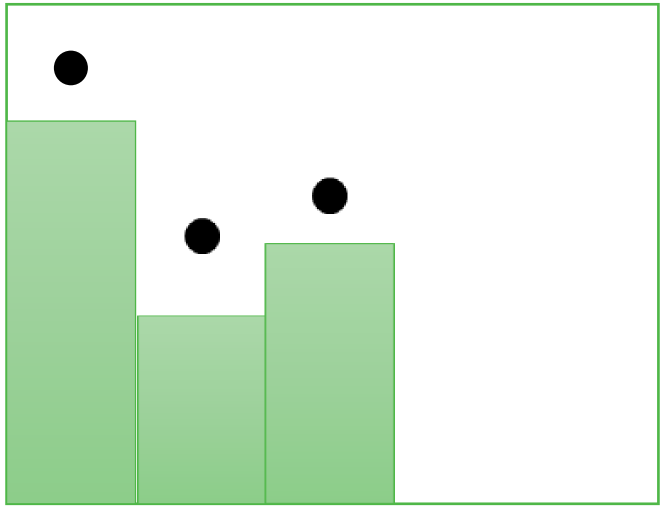


Top quark backgrounds



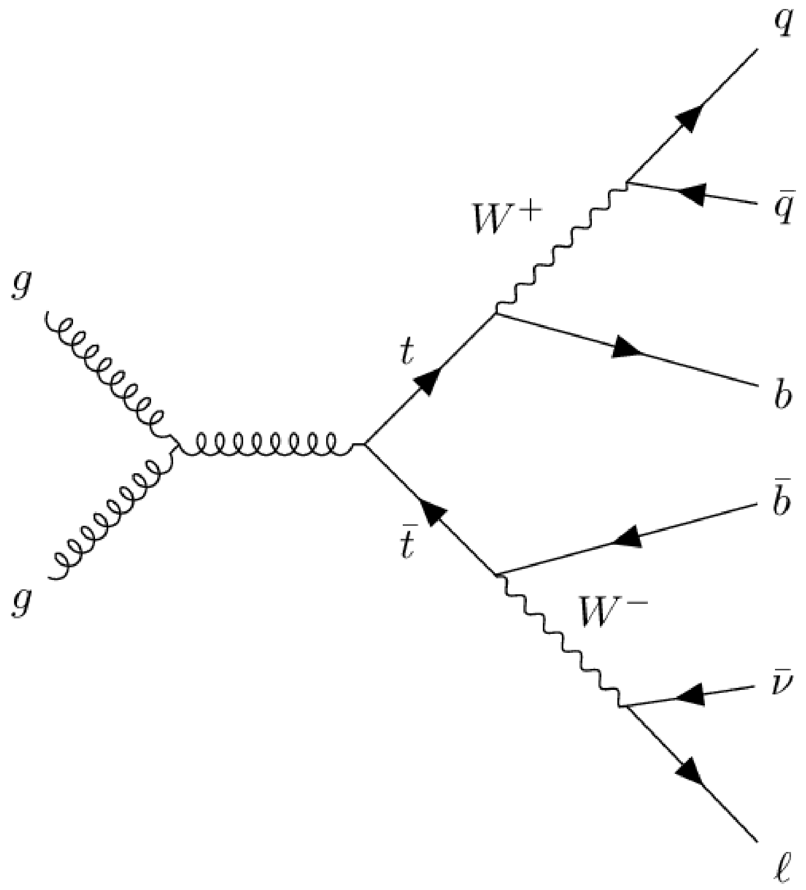
Top quark backgrounds

- Observed mismodelling of the top pair and single top-production
- Control region where we consider dileptons => This region is completely orthogonal to the SR
- CR is over 99% pure in top pair production events



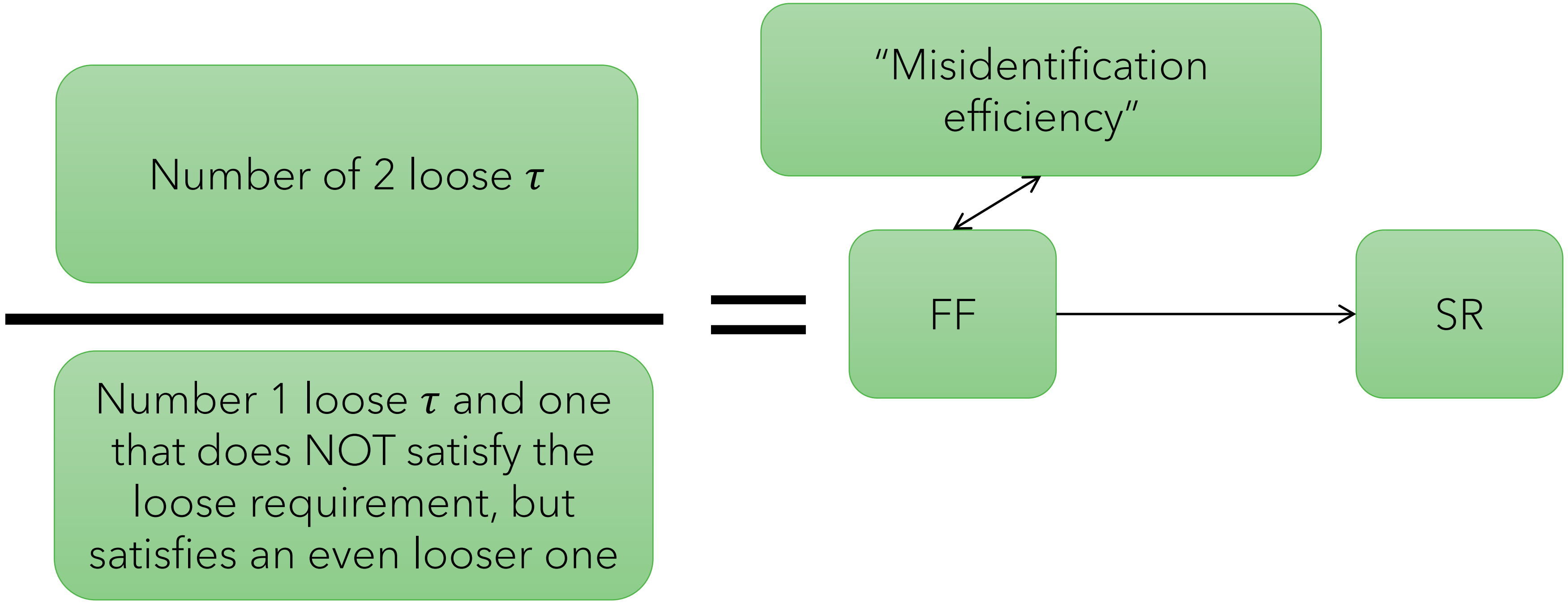
Top backgrounds with jets misidentified as $\tau_{had-vis}$

- Sometimes jets can be misidentified as $\tau_{\mathrm{had-vis}}$, which needs to be accounted for
- CR, except the $\tau_{\mathrm{had-vis}}$ $p_T > 100$ GeV is removed and now $s_T \in [400 - 600]$ GeV
- 97% pure in top pair events, with mixture of correctly identified and misidentified $\tau_{\mathrm{had-vis}}$ (varying with p_T)
- Differentiate between true and misidentified $\tau_{\mathrm{had-vis}}$ using 'transverse mass' parameter -> less neutrinos for jets coming from top pair production



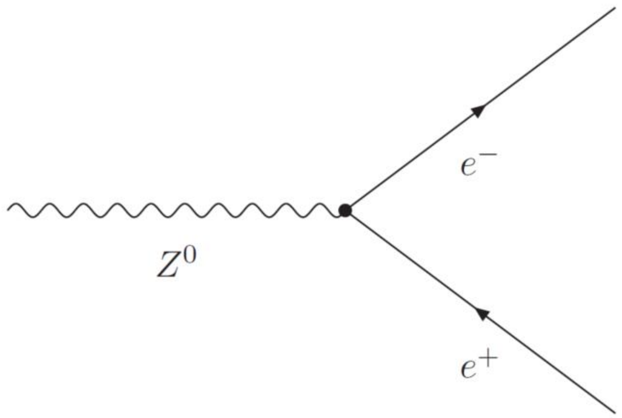
Multi-jet backgrounds with jets misidentified as $\tau_{had-vis}$

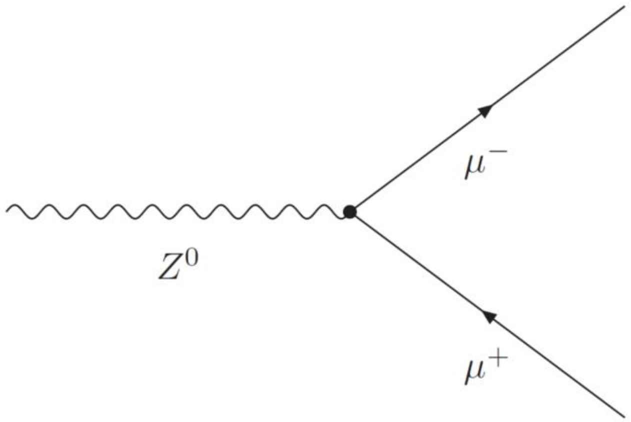
- This applies only to the full hadronic channel
- Few such events -> hard to estimate
- CR -> the two $\tau_{\text{had}-\text{vis}}$ are with same charge, E_T^{miss} loosened



Z+HF background

- There has been observed a small deviation of the cross-section for Z+HF events with theory, therefore data is used to determine the uncertainties
- In the CR the Z boson decays into a light lepton pair (instead of tau) and two heavy-flavor jets
- CR is around 60% Z+HF and 40% $t\bar{t}$ events ($< 1\%$ from backgrounds with misidentified $\tau_{\text{had-vis}}$), RF for $t\bar{t}$ is included

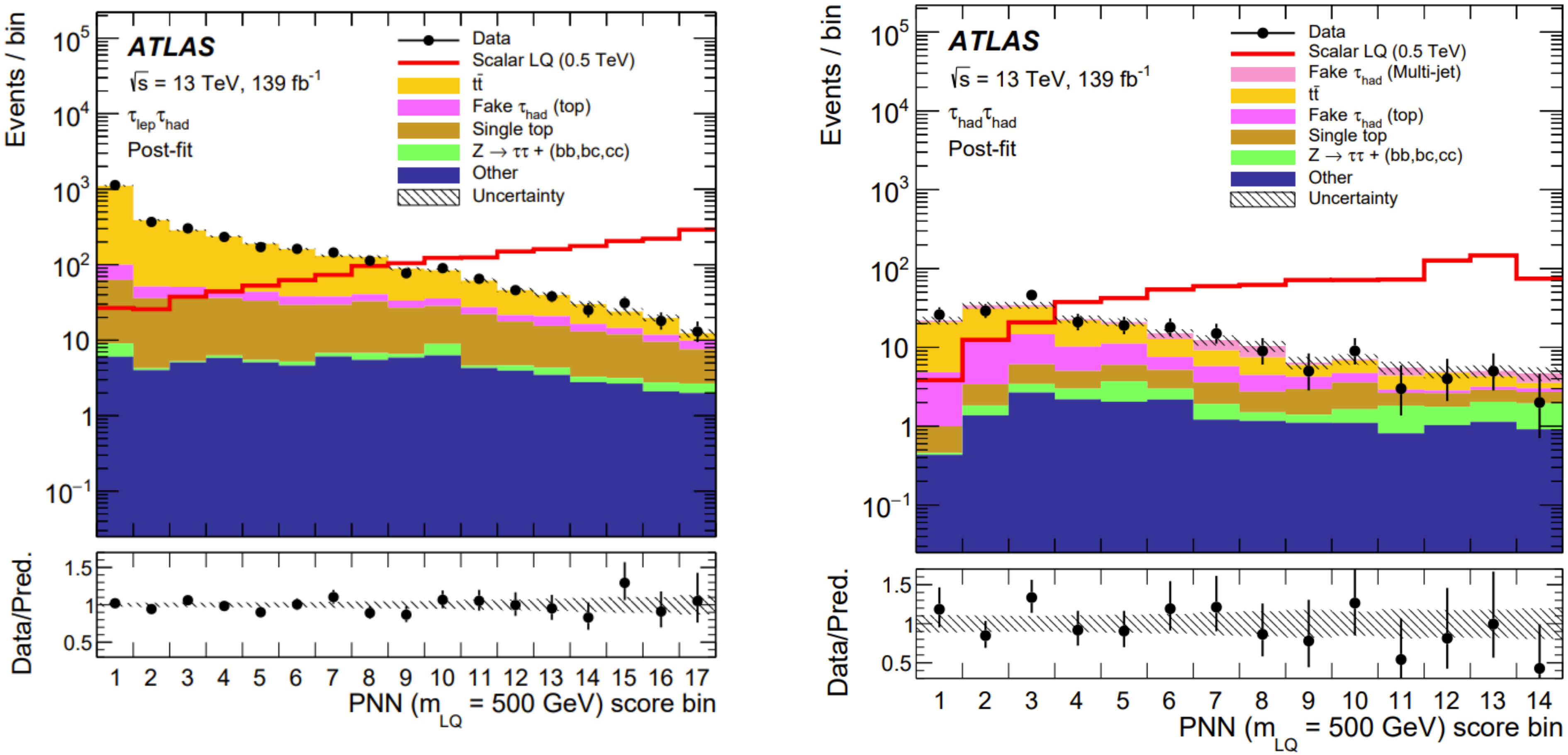




Backgrounds summary

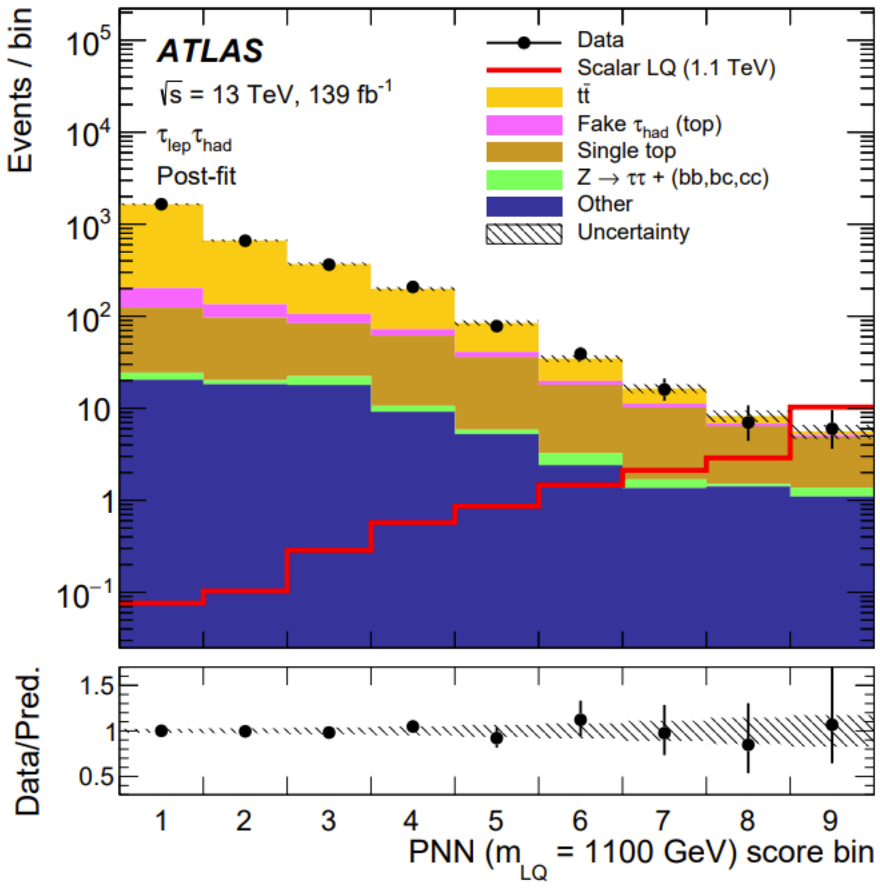
	$\tau_{\text{lep}}\tau_{\text{had}}$ channel	$\tau_{\text{had}}\tau_{\text{had}}$ channel
$t\bar{t}$	2420 ± 90	93 ± 9
Single-top	355 ± 27	20 ± 4
Fake τ_{had} (top)	170 ± 90	43 ± 18
$Z \rightarrow \tau\tau + (bb, bc, cc)$	13.9 ± 2.4	10.3 ± 1.4
Multi-jet	-	22 ± 11
Other	78 ± 7	19 ± 5
Total background	3040 ± 60	207 ± 13
Data	3031	211

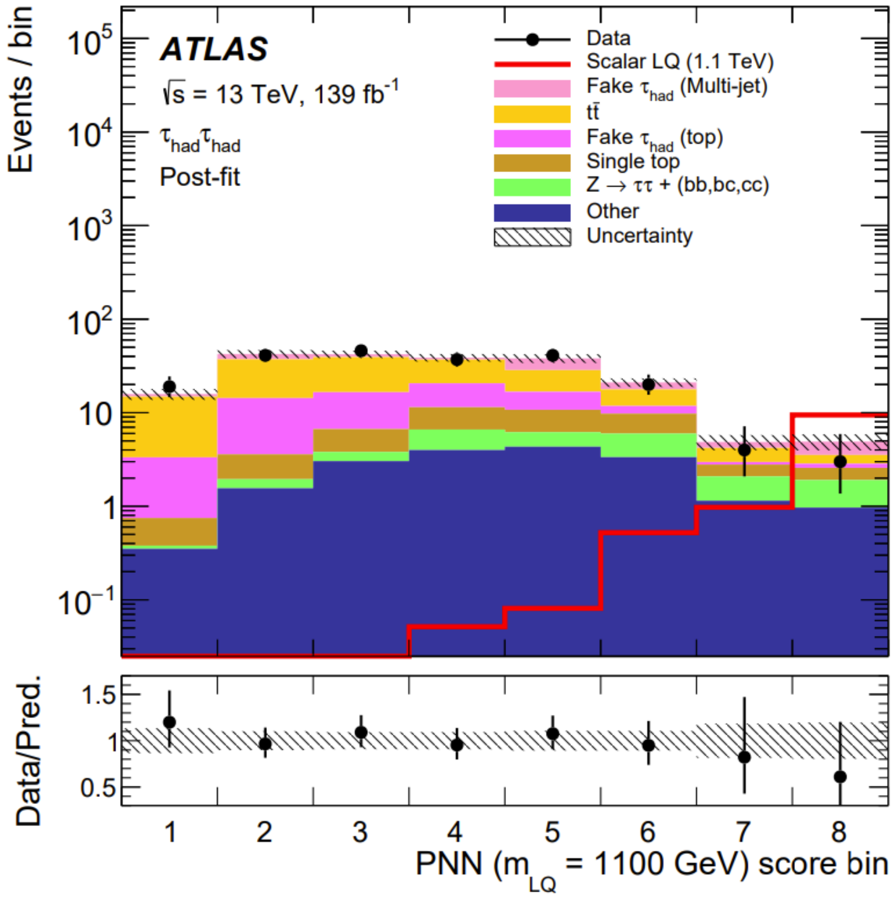
Results



PNN score distributions for $m_{LQ} = 500$ GeV in the semileptonic channel (left) and full hadronic channel (right). Expected signals for the LQ are overlaid.

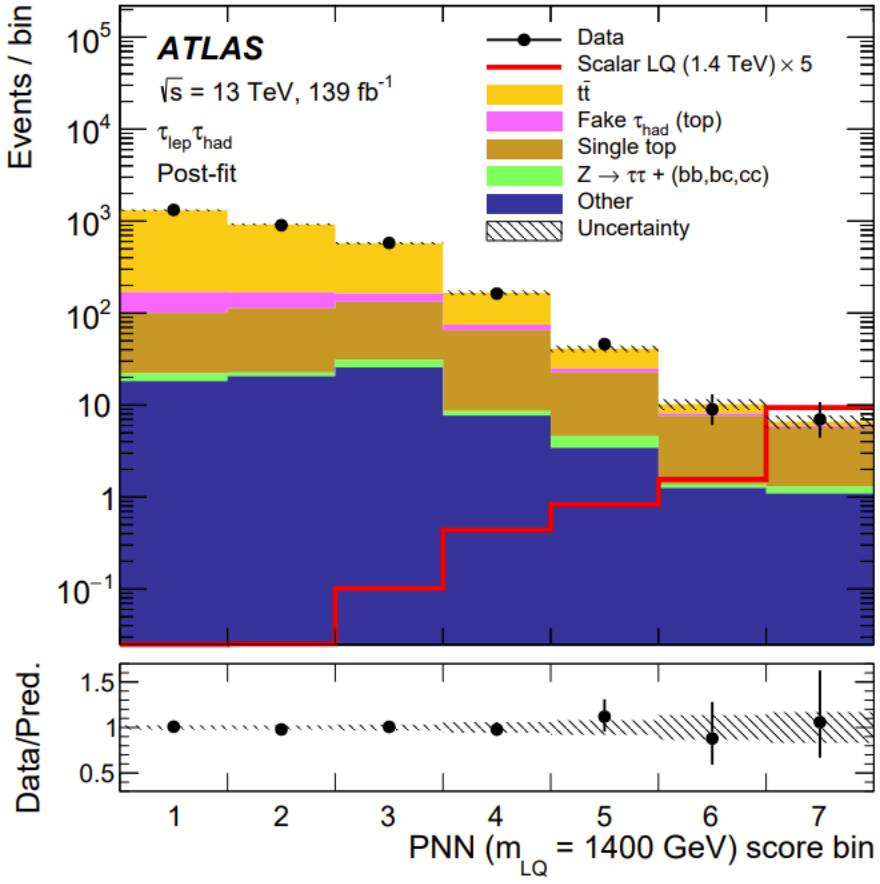
Results

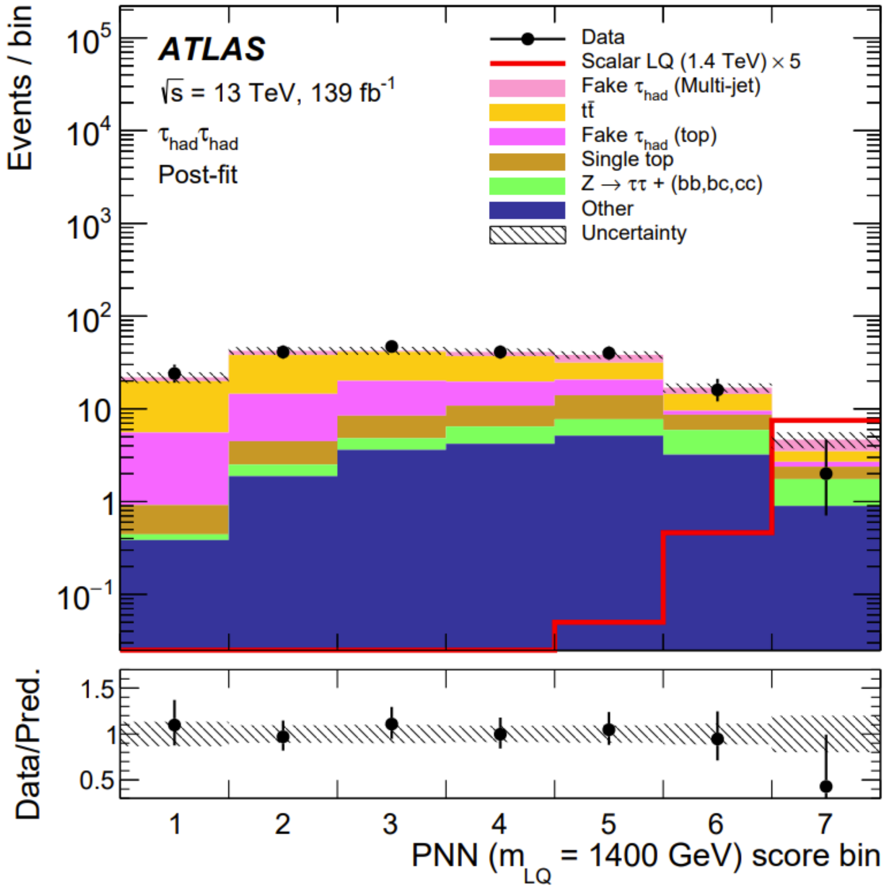




PNN score distributions for $m_{LQ} = 1100$ GeV in the semileptonic channel (left) and full hadronic channel (right). Expected signals for the LQ are overlaid.

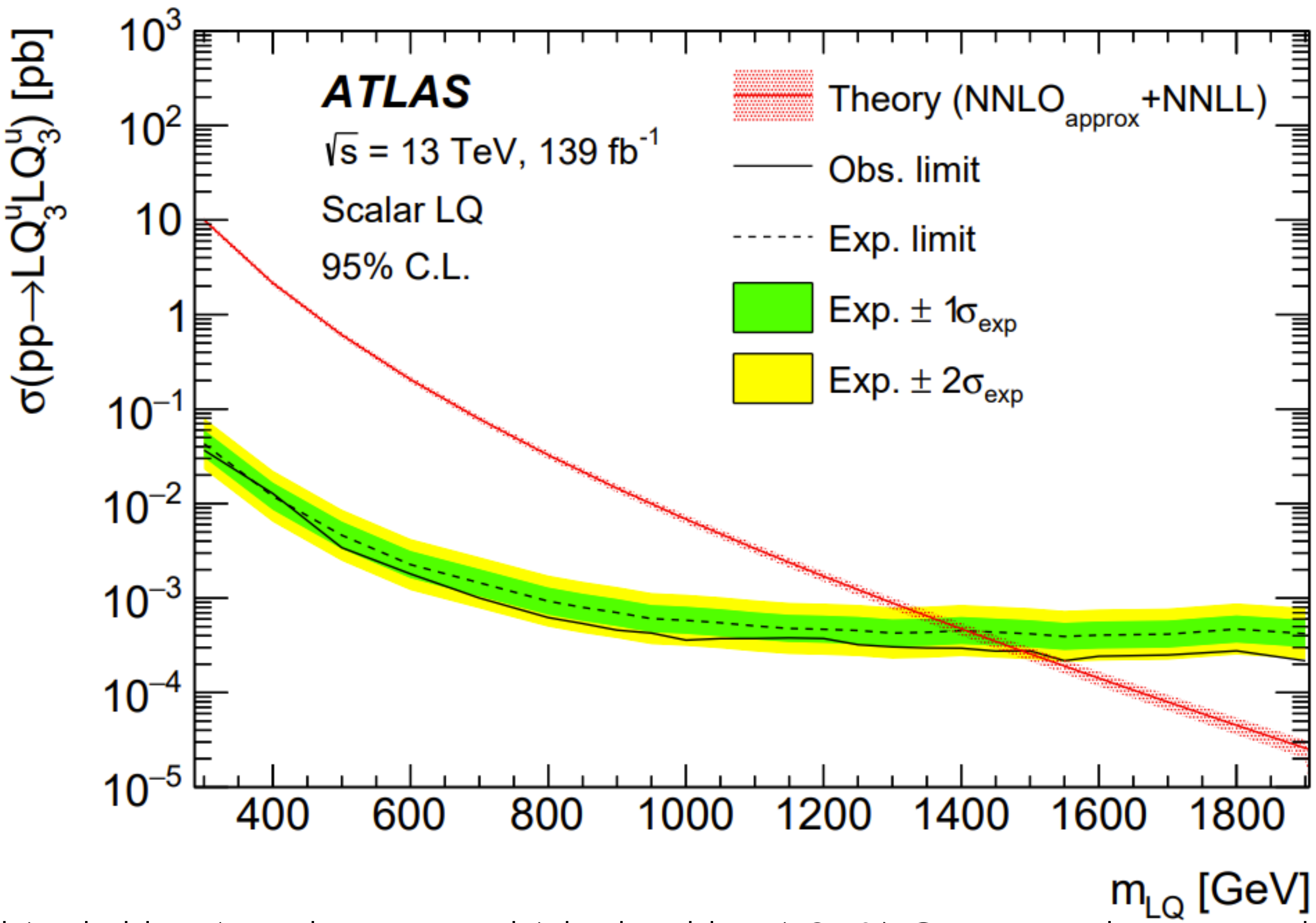
Results





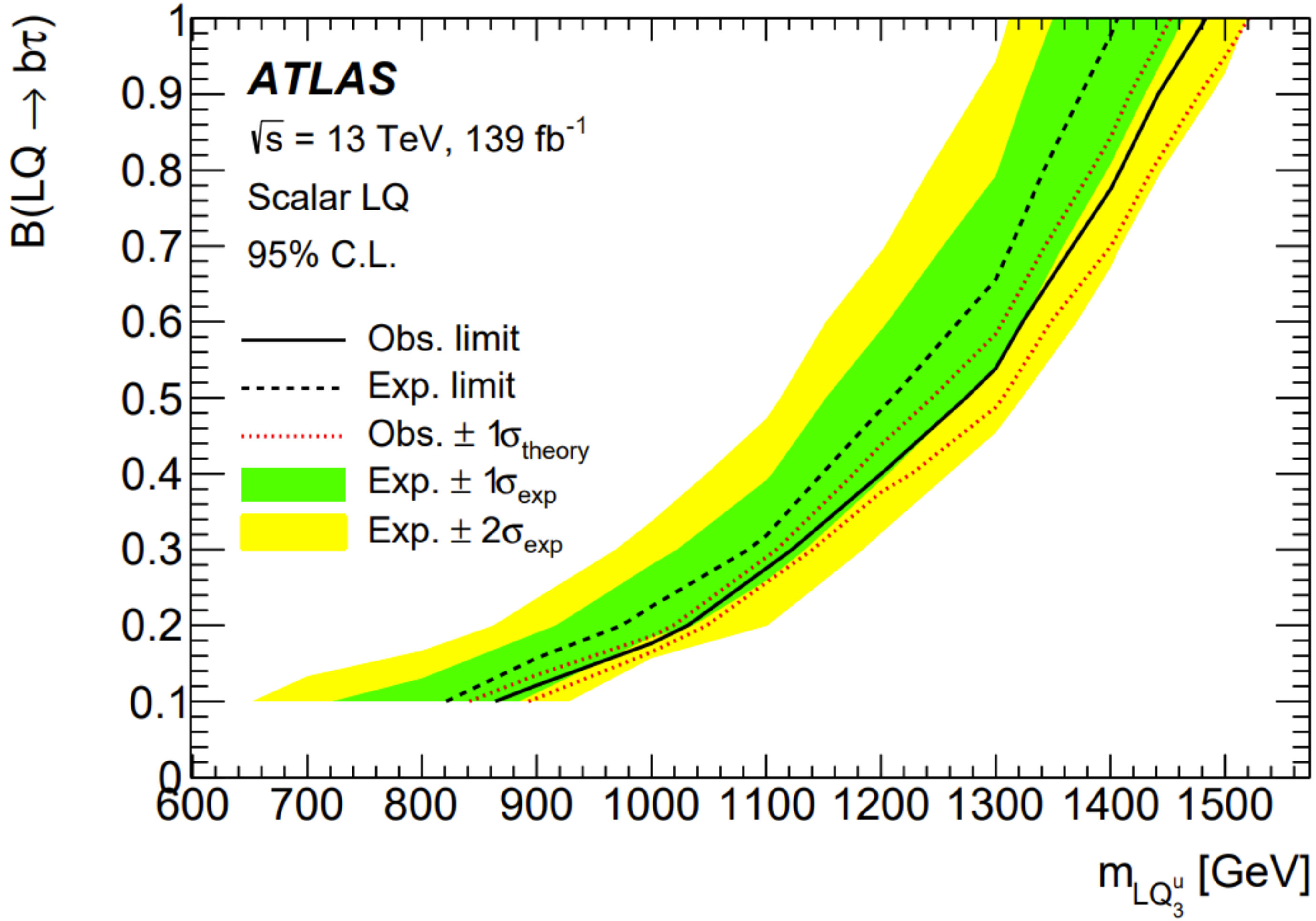
PNN score distributions for $m_{LQ} = 1400$ GeV in the semileptonic channel (left) and full hadronic channel (right). Expected signals for the LQ are overlaid.

Results



Observed (solid line) and expected (dashed line) 95% CL upper limits on the LQ pair production cross section assuming $\mathfrak{B} = 1$ as a function of m_{LQ} in the scalar LQ case. The ± 1 and ± 2 standard deviations from the expected values (green and yellow) are also shown.

Results



Observed (solid line) and expected (dashed line) 95% CL upper limits on the branching ratio as a function of m_{LQ} in the scalar LQ case. The ± 1 and ± 2 standard deviations from the expected values (green and yellow) are also shown.

Results

Possible LQ masses	Scalar LQ	MC LQ	YM LQ
$\mathcal{B} = 1$	> 1490 GeV	> 1690 GeV	> 1960 GeV
$\mathcal{B} = 0.1$	> 850 GeV	> 1100 GeV	> 1300 GeV

- Results here are in 95% CL
- We see no huge dropoff for smaller branching ratios
- Overall, background is in agreement with the SM prediction

Conclusion

- Parametrized Neural Network using the mass as an input gives smooth and precise results
- Precise background estimation with carefully thought out CRs, RFs, SFs, FFs
- Improved estimation of the LQ mass by around 200 GeV for vector LQs and 450 GeV for scalar LQs
- No significant deviations from the SM observed so far
- One could hope to look for LQs in higher-energy regions in the future

Thank You for the attention!

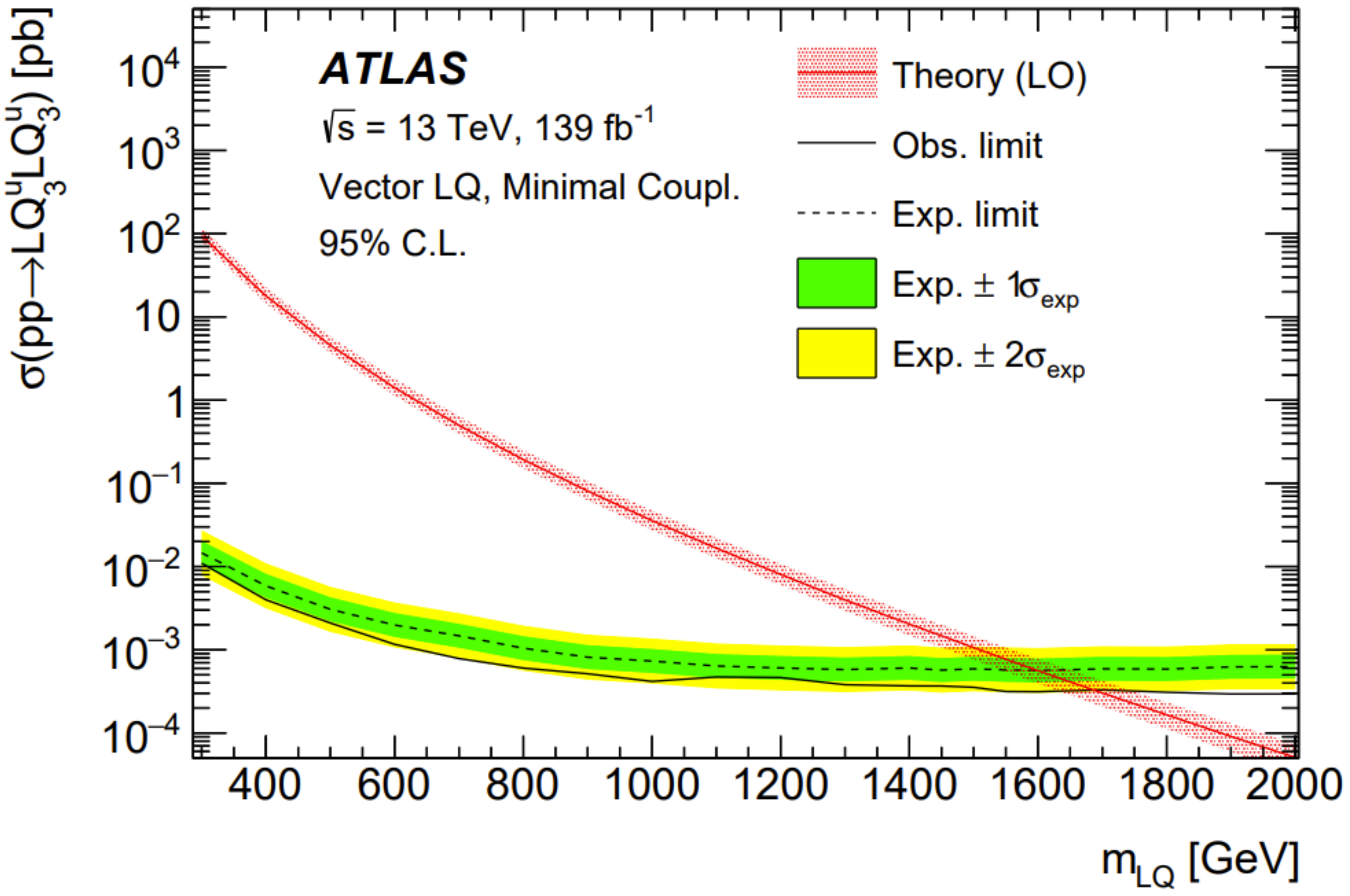
QUESTIONS?

BACKUP SLIDES

Systematic uncertainties

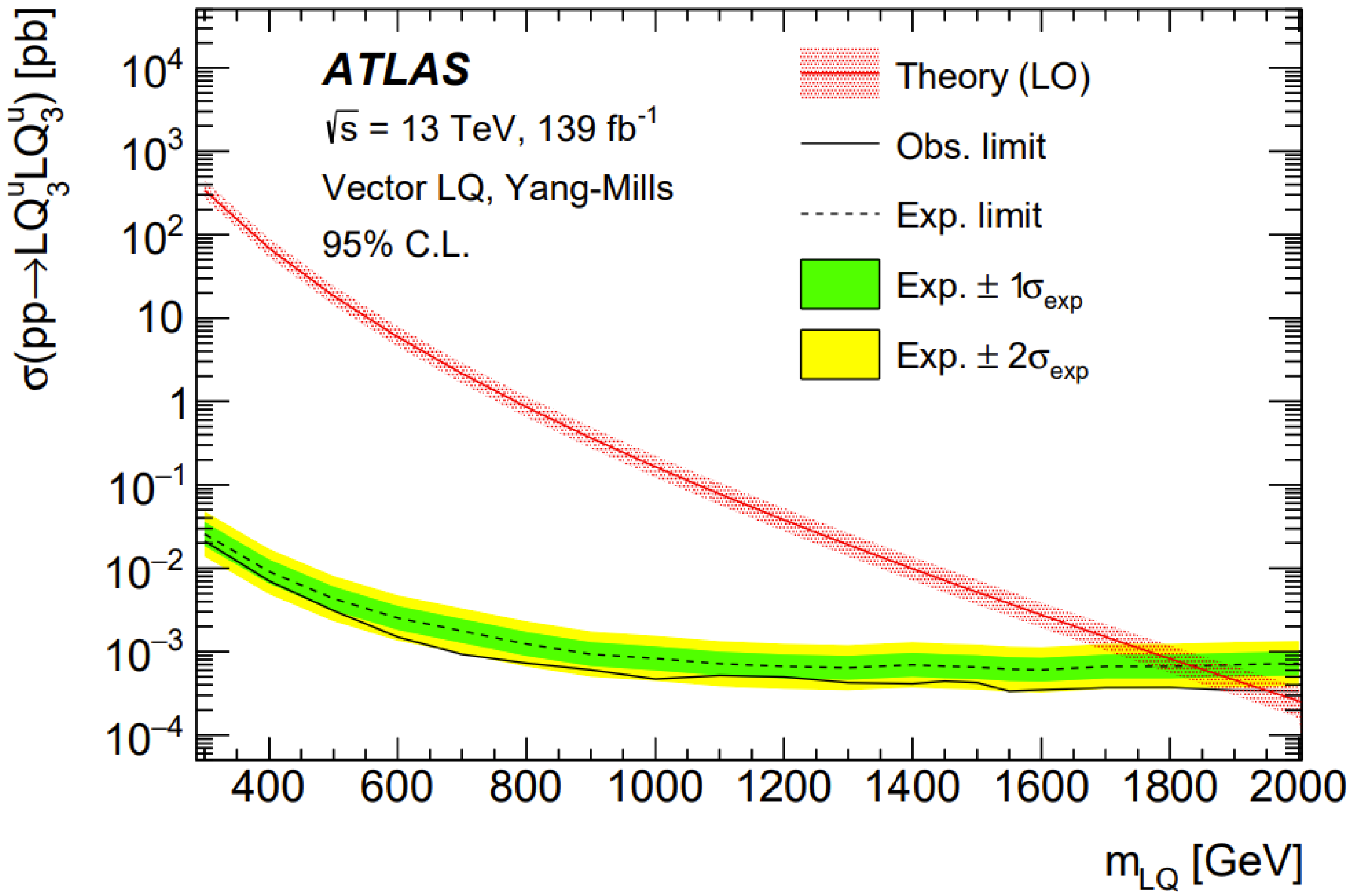
- SM modelling uncertainties (simulations and theoretical frameworks)
- Detector related uncertainties, e.g. signal acceptance
- Main uncertainties come from the top-pair and single top-quark modelling uncertainties
- Correlation between the systematic uncertainties among the SR

Results



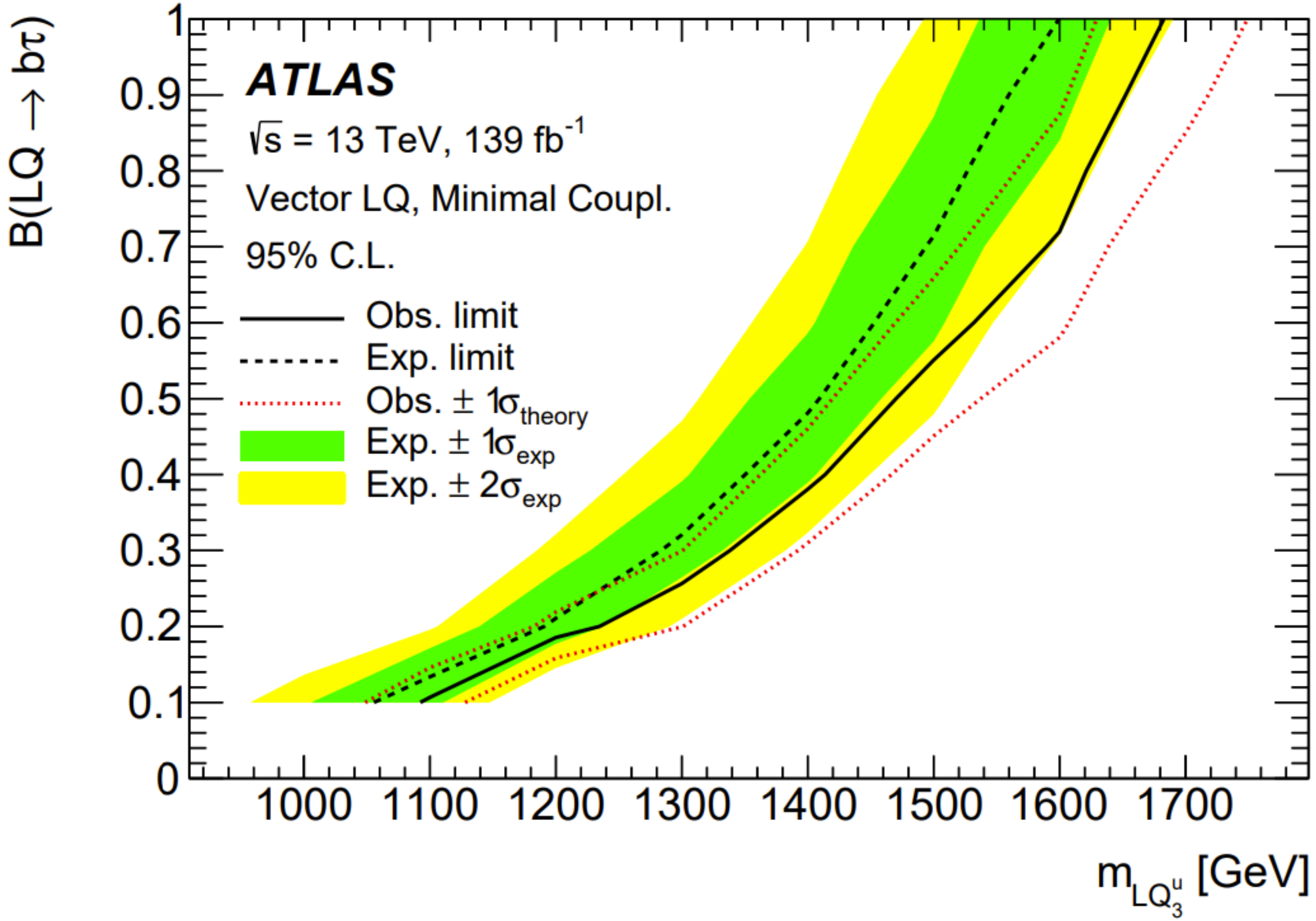
Observed (solid line) and expected (dashed line) 95% CL upper limits on the LQ pair production cross section assuming $\mathfrak{B} = 1$ as a function of m_{LQ} in the vector LQ (Minimal Coupling) case. The ± 1 and ± 2 standard deviations from the expected values (green and yellow) are also shown.

Results



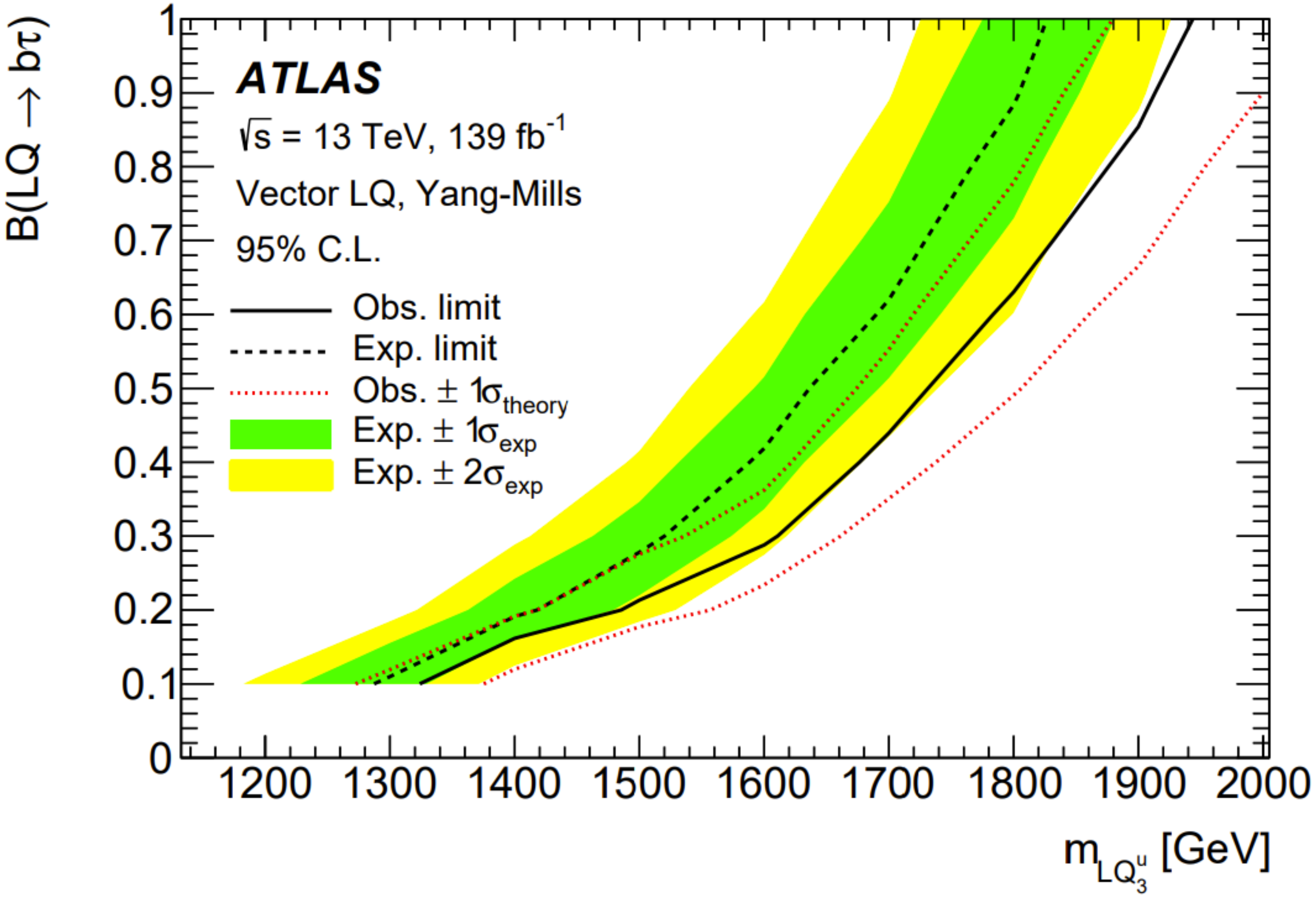
Observed (solid line) and expected (dashed line) 95% CL upper limits on the LQ pair production cross section assuming $\mathfrak{B} = 1$ as a function of m_{LQ} in the vector LQ (Yang-Mills) case. The ± 1 and ± 2 standard deviations from the expected values (green and yellow) are also shown.

Results



Observed (solid line) and expected (dashed line) 95% CL upper limits on the branching ratio as a function of m_{LQ} in the vector LQ (Minimal Couling) case. The ± 1 and ± 2 standard deviations from the expected values (green and yellow), as well as ± 1 standard deviations of the are also shown.

Results



Observed (solid line) and expected (dashed line) 95% CL upper limits on the branching ratio as a function of m_{LQ} in the vector LQ (Yang-Mills) case. The ± 1 and ± 2 standard deviations from the expected values (green and yellow), as well as ± 1 standard deviations of the are also shown.

Event reconstruction and object definitions

- Neutrinos are ‘detected’ using the missing transverse momentum
- At least one pp interaction vertex needed, reconstructed from two or more charged-particle tracks with $p_T > 500$ MeV, the one with the highest summed squared transverse momentum is considered the primary one

Electron candidates

- $p_T > 7$ GeV, $|\eta| < 2.47$, excluding the region $1.37 < |\eta| < 1.52$
- "veto electrons" -> loose identification working point
- "signal electrons" -> tight identification working point

Table 2: Definition of the electron isolation working points and isolation efficiency ϵ. In the Gradient working point definition, the unit of p_{T} is GeV. All working points use a cone size of $\Delta R = 0.2$ for calorimeter isolation and $\Delta R_{\mathrm{max}} = 0.2$ for track isolation.

Working point	Calorimeter isolation	Track isolation
Gradient	$\epsilon = 0.1143 \times p_{\mathrm{T}} + 92.14\%$ (with $E_{\mathrm{T}}^{\mathrm{cone20}}$)	$\epsilon = 0.1143 \times p_{\mathrm{T}} + 92.14\%$ (with $p_{\mathrm{T}}^{\mathrm{varcone20}}$)
HighPtCaloOnly	$E_{\mathrm{T}}^{\mathrm{cone20}} < \max(0.015 \times p_{\mathrm{T}}, 3.5\ \mathrm{GeV})$	-
Loose	$E_{\mathrm{T}}^{\mathrm{cone20}}/p_{\mathrm{T}} < 0.20$	$p_{\mathrm{T}}^{\mathrm{varcone20}}/p_{\mathrm{T}} < 0.15$
Tight	$E_{\mathrm{T}}^{\mathrm{cone20}}/p_{\mathrm{T}} < 0.06$	$p_{\mathrm{T}}^{\mathrm{varcone20}}/p_{\mathrm{T}} < 0.06$

Muon candidates

- $p_T > 7$ GeV, $|\eta| < 2.7$
- “veto muons” – loose identification working point, includes $|\eta| < 0.1$
- “signal muons” -> “medium/high-pT” working point if pT is less/greater than 800 GeV
- The more stringent high-pT requirements remove around 20% of muons but improve resolution by approx. 30% for >1.5 TeV

Muon working points

Table 1: Prompt-muon efficiencies ϵ_μ and light-hadron misidentification rates ϵ_{had} for the different selection working points, evaluated in a $t\bar{t}$ MC sample in different p_T regions for $|\eta| < 2.5$. It should be noted that the *Tight* WP by construction does not select any muons with $p_\text{T} < 4$ GeV, which is reflected in the corresponding efficiency in the first p_T region. The statistical uncertainties are at least one order of magnitude smaller than the last digit reported.

	$3 < p_\text{T}$ [GeV] < 5		$5 < p_\text{T}$ [GeV] < 20		$20 < p_\text{T}$ [GeV] < 100		$p_\text{T} > 100$ GeV	
Selection WP	ϵ_μ [%]	ϵ_{had} [%]	ϵ_μ [%]	ϵ_{had} [%]	ϵ_μ [%]	ϵ_{had} [%]	ϵ_μ [%]	ϵ_{had} [%]
Loose	90	1.17	98	1.06	99	0.25	98	0.12
Medium	70	0.63	97	0.85	97	0.17	97	0.07
Tight	36	0.15	90	0.38	93	0.12	93	0.04
Low-p_T (cut-based)	86	0.82	95	0.71	97	0.17	97	0.07
Low-p_T (multivariate)	88	0.73	96	0.66	97	0.17	97	0.07
High-p_T	45	0.34	79	0.60	80	0.13	80	0.05

Muon working points

Table 3: Isolation efficiencies for prompt muons, ϵ_μ, and muons from bottom and charm semileptonic decays, ϵ_{HF}, for the different isolation working points, evaluated in a $t\bar{t}$ MC sample in different p_T regions for tracks satisfying the *Medium* identification and the vertex association criteria. The isolation working points considered correspond to the variants with the cone size remaining constant at $\Delta R = 0.2$ for $p_T^\mu > 50$ GeV. The statistical uncertainties are at least one order of magnitude smaller than the last digit reported.

	$3 < p_T$ [GeV] < 5		$5 < p_T$ [GeV] < 20		$20 < p_T$ [GeV] < 100		$p_T > 100$ GeV	
Isolation WP	ϵ_μ [%]	ϵ_{HF} [%]	ϵ_μ [%]	ϵ_{HF} [%]	ϵ_μ [%]	ϵ_{HF} [%]	ϵ_μ [%]	ϵ_{HF} [%]
Loose	63	14.3	86	7.2	97	6.1	99	12.7
Tight	53	11.9	70	4.2	89	1.0	98	1.6
PflowLoose	62	12.9	86	6.8	97	5.0	99	9.1
PflowTight	45	8.5	63	3.1	87	0.9	97	0.8
HighPtTrackOnly	92	35.9	92	17.2	92	4.5	92	0.6
TightTrackOnly	80	19.9	81	7.0	94	3.2	99	3.3
PLBDTLoose	81	17.4	83	5.1	93	1.3	98	1.7
PLBDTTight	57	9.6	69	2.7	87	0.5	98	1.7

Jet reconstruction

- Jets are reconstructed from topological energy clusters, using the anti-k_t algorithm

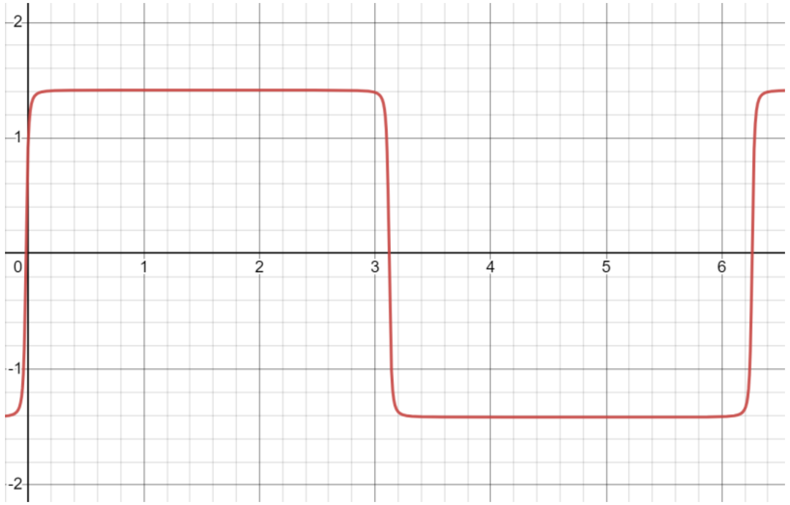
$$d_{ij} = \min(k_{ti}^{2p}, k_{tj}^{2p}) \frac{\Delta_{ij}^2}{R^2}$$
$$d_{iB} = k_{ti}^{2p}$$

- Here $\Delta_{ij}^2 = \left(y_i - y_j\right)^2 + (\phi_i - \phi_j)^2$ and k is the transverse momentum, y the rapidity and ϕ the azimuth of the particle
- In the anti-k_t algorithm, one uses $p = -1$
- The algorithm identifies the smallest of the distances, recombining entities or removing them as such if d_{iB} is the smallest

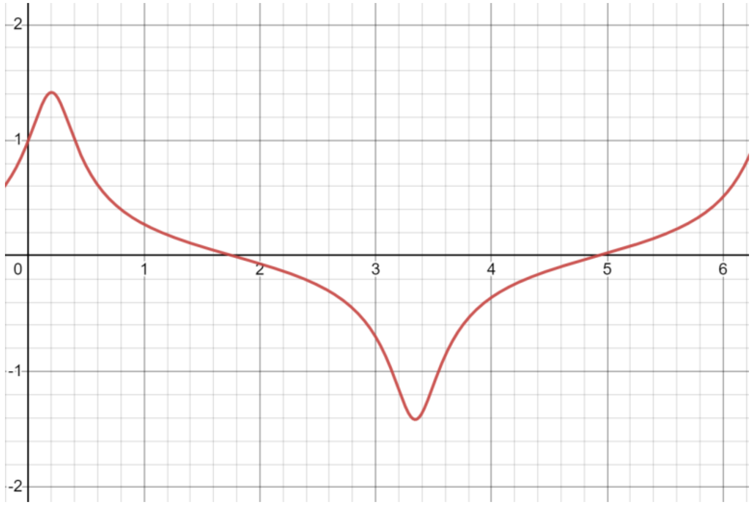
$p_T^{miss}\phi$ centrality

$$\frac{A+B}{\sqrt{A^2+B^2}}$$

$$A = \frac{\sin(\phi_{p_T^{miss}} - \phi_{\tau_2})}{\sin(\phi_{\tau_1} - \phi_{\tau_2})}; B = \frac{\sin(\phi_{\tau_1} - \phi_{p_T^{miss}})}{\sin(\phi_{\tau_1} - \phi_{\tau_2})}$$



Two products in almost opposite directions



The two product with relatively small azimuthal opening angle between them

Additional ambiguities resolving technique

- First, electron candidates are discarded if they share a track with a more energetic electron or a muon identified in the MS; if the muon is identified in the calorimeter it is removed instead. Any τhad-vis candidate within ΔR = 0.2 of an electron or a muon (which must be reconstructed in the MS if the τhad-vis pT is above 50 GeV) is then rejected. Jets are discarded if they lie within ΔR = 0.2 of an electron or have fewer than three associated tracks and lie within the same distance of a muon. Electron or muon (τhad-vis) candidates within ΔR = 0.4 (ΔR = 0.2) of any remaining jet are then removed. Finally, ambiguities between anti-τhad-vis candidates and jets within ΔR = 0.2 are resolved in favour of the jet if it is b-tagged or the anti-τhad-vis otherwise

Additional info about overall top bckgd CR

- Two b-jets with $p_T > 45, 20$ GeV
- Two light leptons with opposite charges
- $E_T^{miss} > 100$ GeV
- Dilepton mass $m_{\ell\ell} > 110$ GeV
- $m_{b\ell} > 250$ GeV, where $m_{b\ell} = \min(\max(m_{b0\ell0}, m_{b1\ell1}), \max(m_{b0\ell1}, m_{b1\ell0}))$ with (0) denoting leading and (1) sub-leading

Info about the simulation and data samples

- Integrated luminosity of $139\ fb^{-1}$ for the collision data collected between 2015 and 2018
- An average of $< \mu >$ additional interactions equal to $\mathbf{33.7}$
- For LQ masses between $300\ GeV$ and $2000\ GeV$ the cross-sections range from 10 pb to 0.01 fb

MMC

- 6-8 unknowns, only 4 equations
- Some solutions are more likely than others which can let us do a better estimation

$$\not{E}_{\mathrm{T}_x} = p_{\mathrm{mis}_1} \sin\theta_{\mathrm{mis}_1} \cos\phi_{\mathrm{mis}_1} + p_{\mathrm{mis}_2} \sin\theta_{\mathrm{mis}_2} \cos\phi_{\mathrm{mis}_2}$$

$$\not{E}_{\mathrm{T}_y} = p_{\mathrm{mis}_1} \sin\theta_{\mathrm{mis}_1} \sin\phi_{\mathrm{mis}_1} + p_{\mathrm{mis}_2} \sin\theta_{\mathrm{mis}_2} \sin\phi_{\mathrm{mis}_2}$$

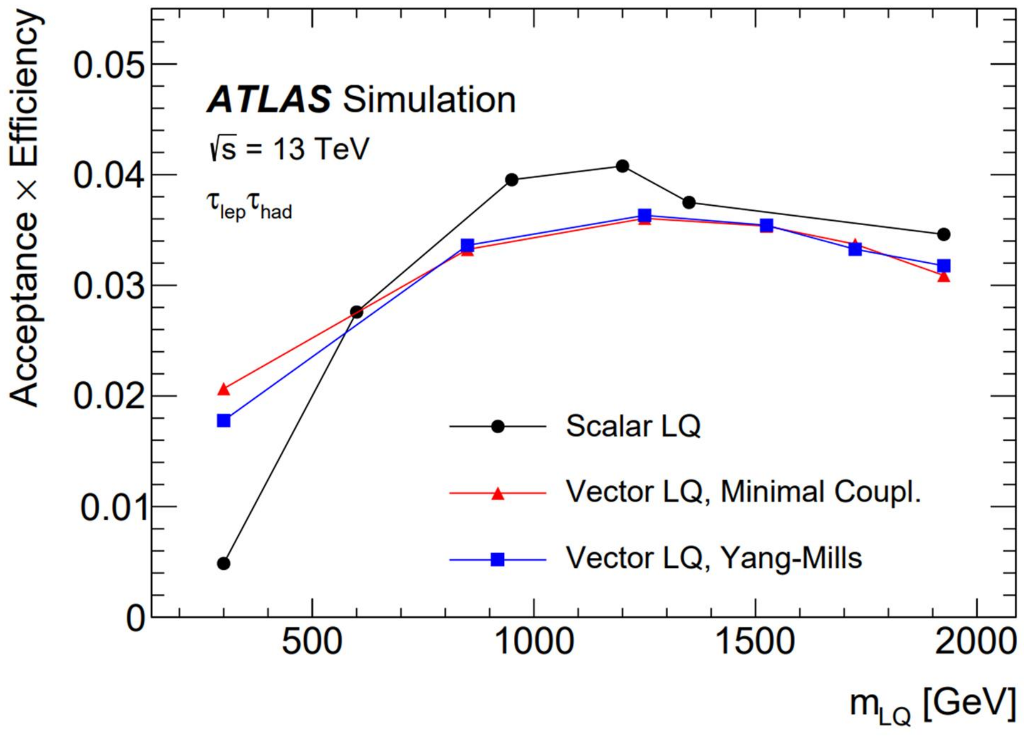
$$M_{\tau_1}^2 = m_{\mathrm{mis}_1}^2 + m_{\mathrm{vis}_1}^2 + 2\sqrt{p_{\mathrm{vis}_1}^2 + m_{\mathrm{vis}_1}^2}\sqrt{p_{\mathrm{mis}_1}^2 + m_{\mathrm{mis}_1}^2} - 2p_{\mathrm{vis}_1} p_{\mathrm{mis}_1} \cos\Delta\theta_{vm_1}$$

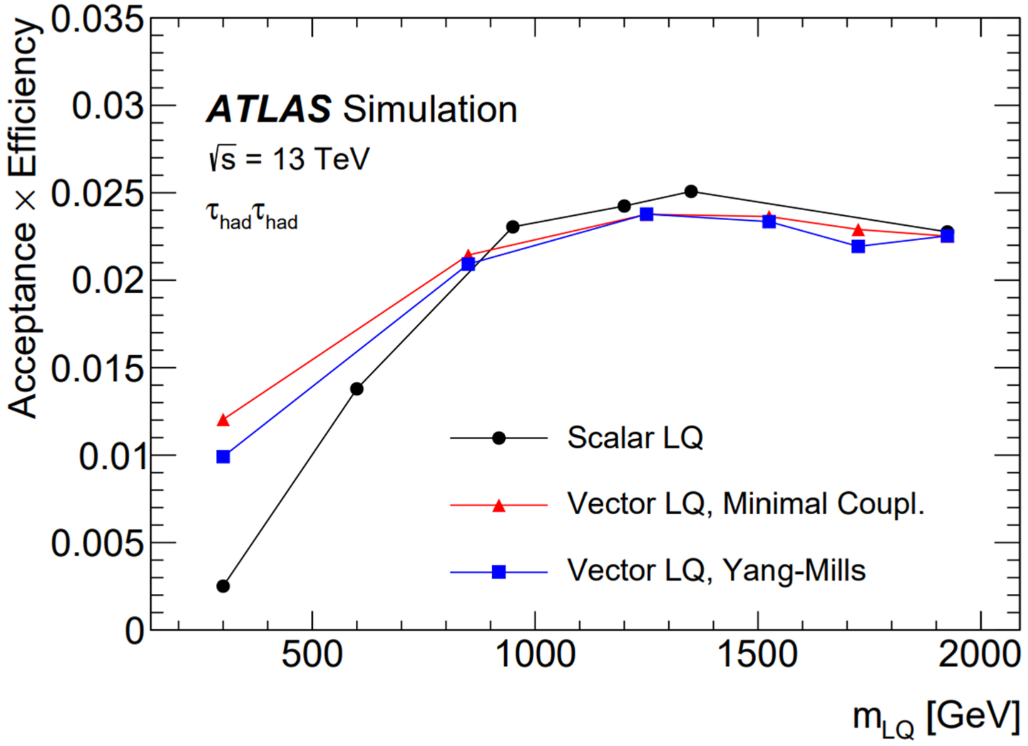
$$M_{\tau_2}^2 = m_{\mathrm{mis}_2}^2 + m_{\mathrm{vis}_2}^2 + 2\sqrt{p_{\mathrm{vis}_2}^2 + m_{\mathrm{vis}_2}^2}\sqrt{p_{\mathrm{mis}_2}^2 + m_{\mathrm{mis}_2}^2} - 2p_{\mathrm{vis}_2} p_{\mathrm{mis}_2} \cos\Delta\theta_{vm_2}$$

Simulation data

Process	ME generator	ME QCD order	ME PDF	PS and hadronisation	UE tune	Cross-section order
Top-quark						
$t\bar{t}^{(\S)}$	Powheg-Box v2 [62]	NLO	NNPDF3.0NLO	Pythia 8.230	A14	NNLO+NNLL [63]
t-channel	Powheg-Box v2	NLO	NNPDF3.0NLO	Pythia 8.230	A14	NLO [64]
s-channel	Powheg-Box v2	NLO	NNPDF3.0NLO	Pythia 8.230	A14	NLO [65]
$Wt^{(\S)}$	Powheg-Box v2	NLO	NNPDF3.0NLO	Pythia 8.230	A14	NLO [66]
Top-quark + W/Z						
$t\bar{t}Z$	Sherpa 2.2.1 [67–69]	NLO	NNPDF3.0NNLO	Sherpa 2.2.1	Default	$\text{NLO}^{(\dagger)}$
$t\bar{t}W$	Sherpa 2.2.8	NLO	NNPDF3.0NNLO	Sherpa 2.2.8	Default	$\text{NLO}^{(\dagger)}$
Vector boson + jets						
W/Z+jets	Sherpa 2.2.1	NLO ($\leq$ 2 jets) LO (3,4 jets)	NNPDF3.0NNLO	Sherpa 2.2.1	Default	NNLO [70]
Diboson						
WW, WZ, ZZ	Sherpa 2.2.1	NLO ($\leq$ 1 jet) LO (2,3 jets)	NNPDF3.0NNLO	Sherpa 2.2.1	Default	$\text{NLO}^{(\dagger)}$
Higgs boson						
ggF	Powheg-Box v2	NNLO	NNPDF3.0NLO	Pythia 8.212	AZNLO [71]	N3LO(QCD)+NLO(EW) [72–76]
VBF	Powheg-Box v2	NLO	NNPDF3.0NLO	Pythia 8.212	AZNLO	NNLO(QCD)+NLO(EW) [72, 77–79]
$qq \to WH$	Powheg-Box v2	NLO	NNPDF3.0NLO	Pythia 8.212	AZNLO	NNLO(QCD)+NLO(EW) [56–59, 61, 80]
$qq \to ZH$	Powheg-Box v2	NLO	NNPDF3.0NLO	Pythia 8.212	AZNLO	$\text{NNLO(QCD)+NLO(EW)}^{(\ddagger)}$
$gg \to ZH$	Powheg-Box v2	NLO	NNPDF3.0NLO	Pythia 8.212	AZNLO	NLO+NLL [81–85]
$t\bar{t}H$	Powheg-Box v2	NLO	NNPDF3.0NLO	Pythia 8.230	A14	NLO [72]

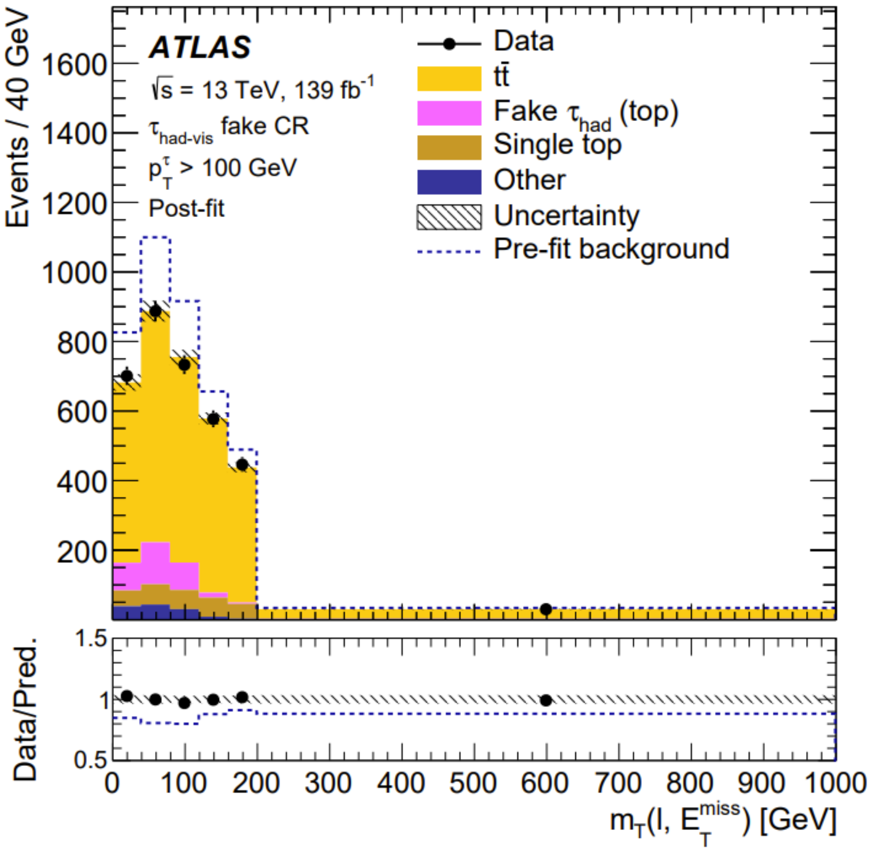
Signal Region

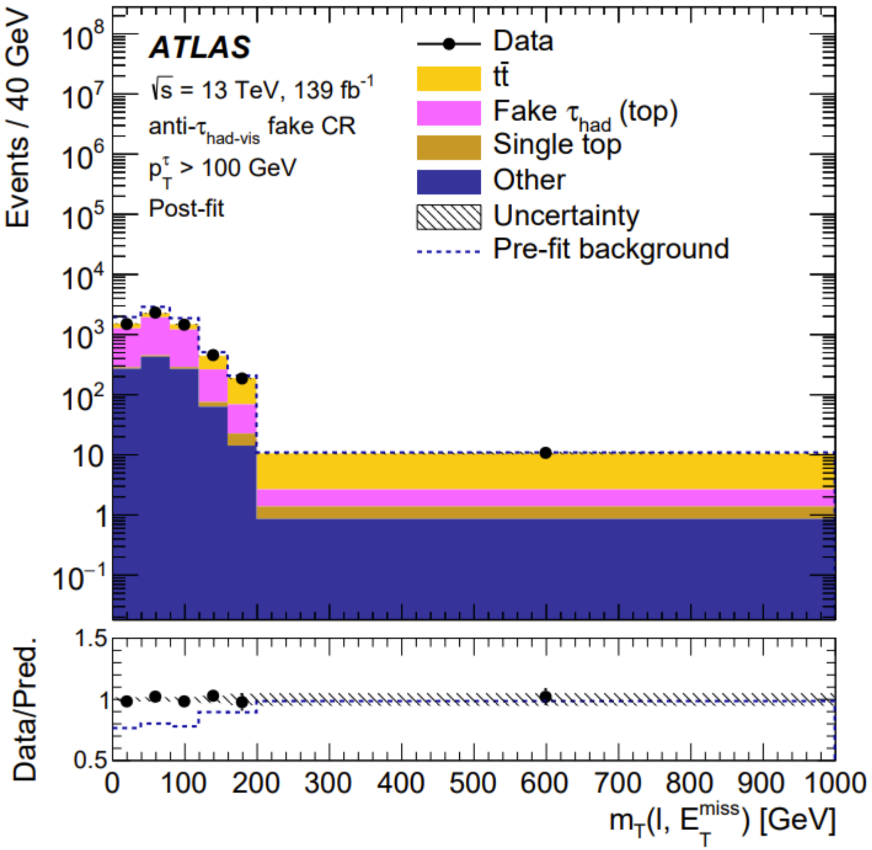




Expected acceptance times efficiency for all the scenarios with $\beta = 0.5$. The $\tau_{\mathrm{lep}}\tau_{\mathrm{had}}$ channel on the left and the $\tau_{\mathrm{had}}\tau_{\mathrm{had}}$ channel is on the right.

Backgrounds with jets misidentified as τ_{had}





Transverse mass definition

$$m_{\mathrm{T}}(\ell, E_{\mathrm{T}}^{\mathrm{miss}}) = \sqrt{(E_{\mathrm{T}}^{\mathrm{miss}} + p_{\mathrm{T},\ell})^2 - (E_{\mathrm{T},x}^{\mathrm{miss}} + p_{x,\ell})^2 - (E_{\mathrm{T},y}^{\mathrm{miss}} + p_{y,\ell})^2}$$