Search for heavy Higgs bosons from a g2HDM in multilepton plus b-jets final states in pp collisions at 13 TeV with the ATLAS detector

Merve Nazlim Agaras on behalf of the team

Red LHC Workshop

11.05.2023

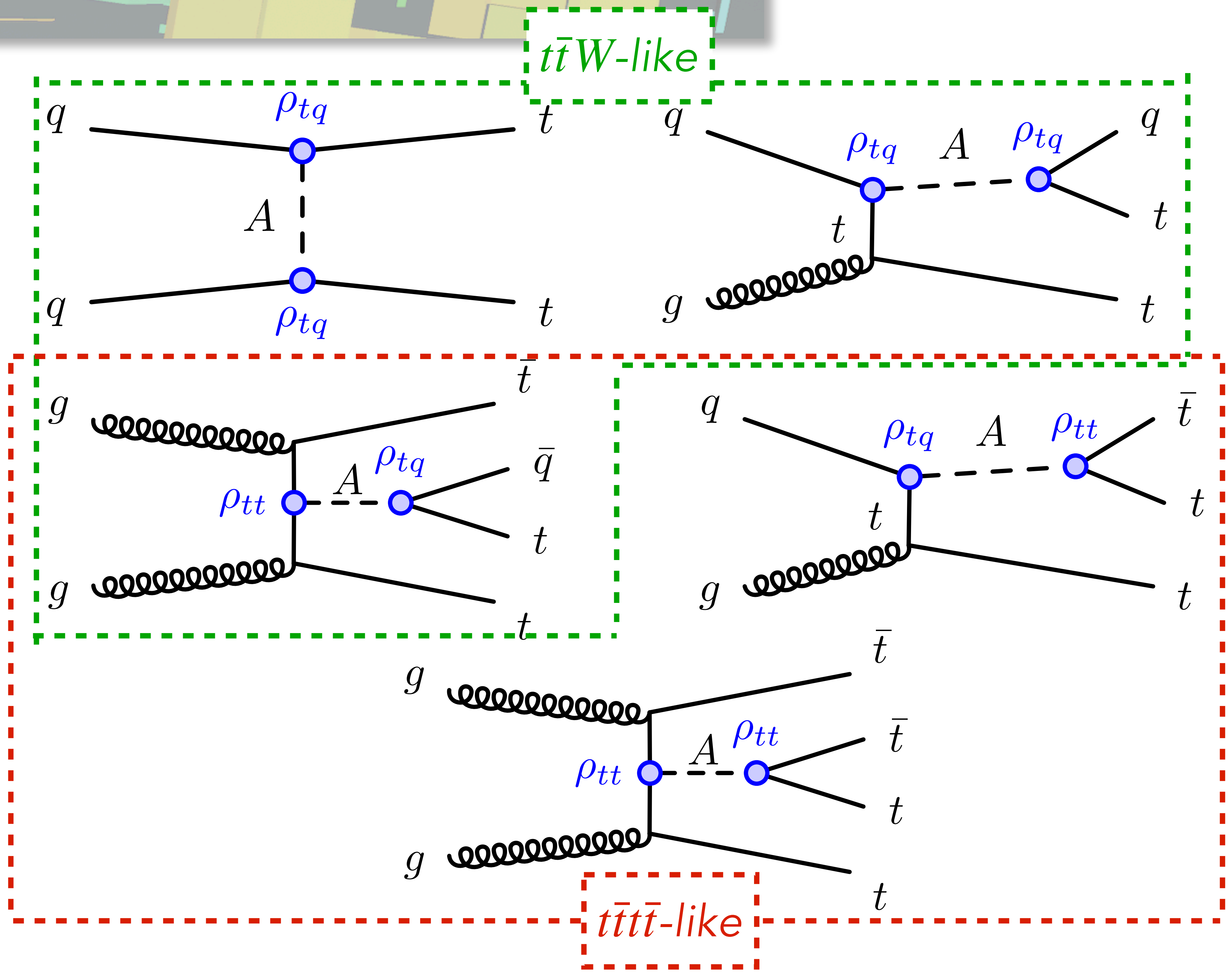
ATLAS EXPERIMENT

IFAE Institut de Física d'Altes Energies

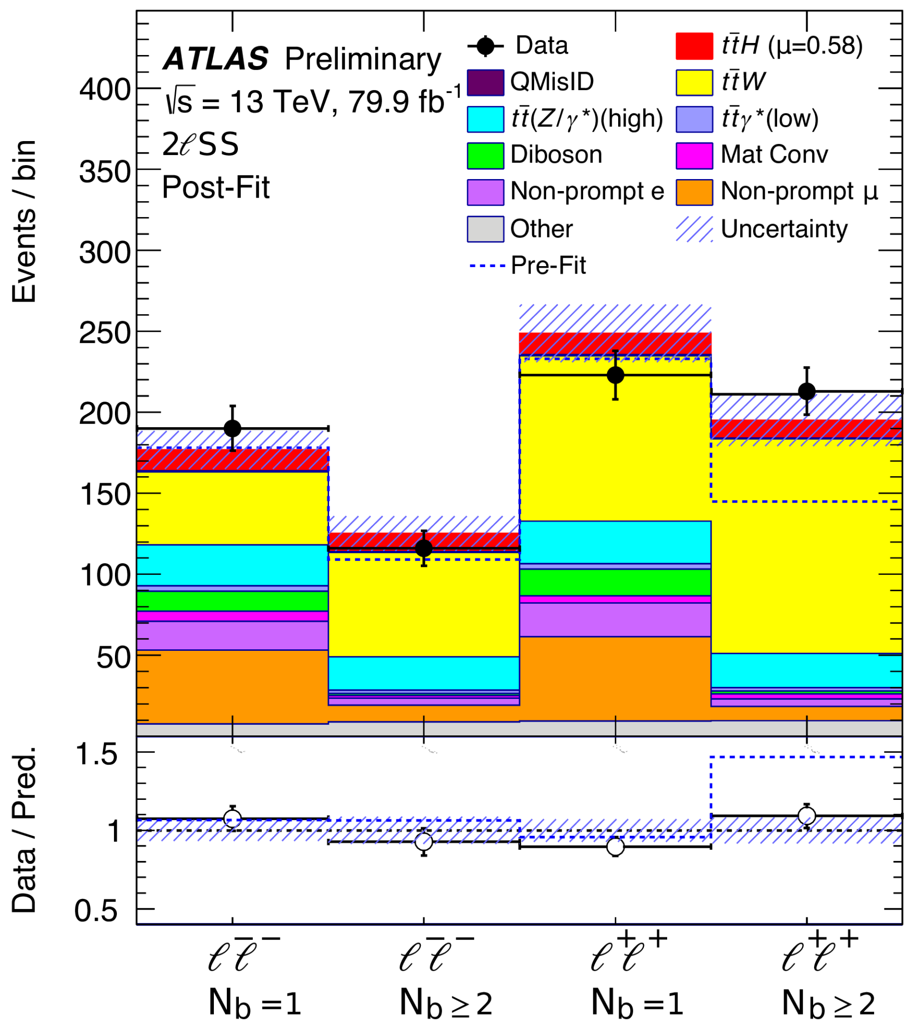
EXCELENCIA SEVERO OCHOA

Motivation

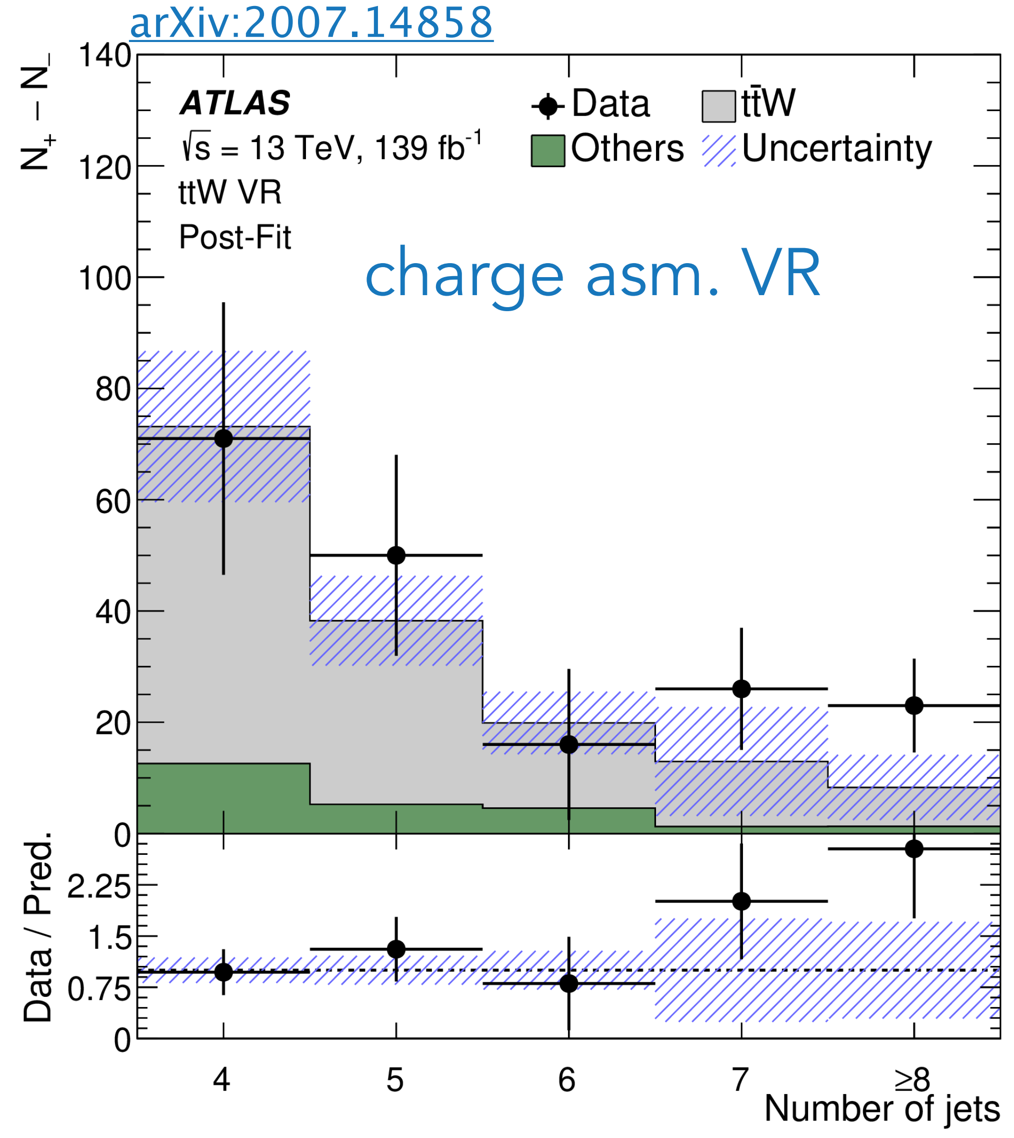
* General 2HDM: including off-diagonal Yukawa couplings for the second doublet in the alignment limit while dropping discrete symmetry (Z_2) which will allow for FCNH
  - Extra **sub-TeV** Higgs bosons (H^0, A^0) with extra Yukawa couplings: ρ_{tt}, ρ_{tu}, ρ_{tc}
  - Satisfy existing experimental constraints (Higgs coupling measurements, flavour physics, etc.), and can address several shortcomings of the SM: electroweak baryogenesis, strong CP problem, flavour problem, etc
  - Production modes with tH and ttH and resulting processes $sstt, ttq, ttt, tttq, tttt$ are studied
  - Final states with multiple leptons (2ℓSS, 3ℓ, 4ℓ) and b-jets studied (expect lepton charge asymmetry)
* Some analyses observed odd data features, that g2HDM model could explain
  - ttHML: tension in 2LSS (++) >=2bj region
  - SM 4top: also observed a lepton charge asymmetry at high Njets
  - Recent 4tops BSM search confirms this deviation and the asymmetric high nJets data/MC discrepancy



ATLAS-CONF-2019-045

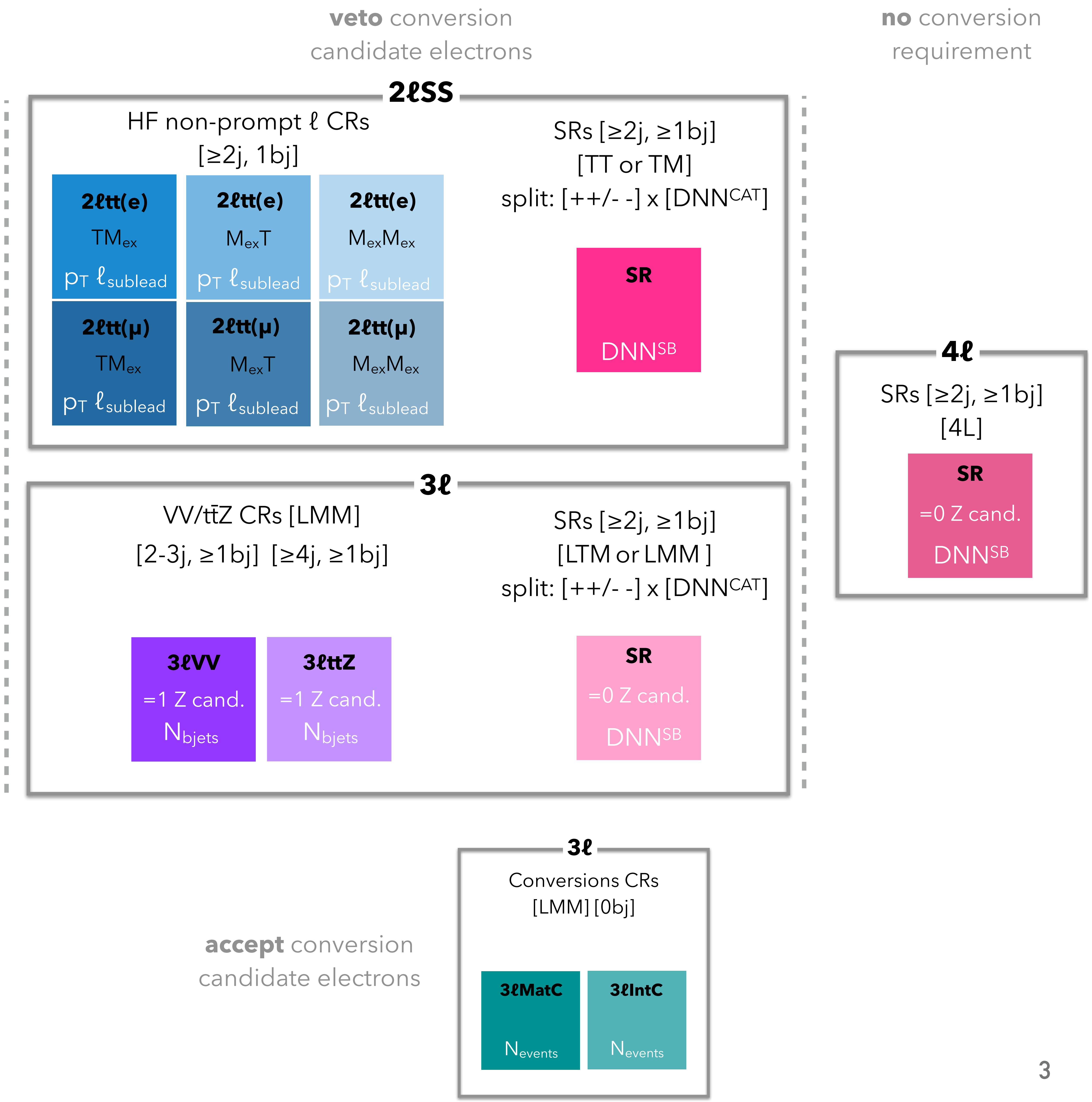


arXiv:2007.14858



Analysis strategy

* Use **tight lepton definition** to suppress reducible background
  - Non-prompt lepton veto BDTs
* **Estimate corrections** from data for various non-prompt ℓ (HF, conversions) and irreducible (VV+HF jets, $t\bar{t}Z$, $t\bar{t}W$) backgrounds in simultaneous fit CRs+SRs
  - Multiple **control regions** are defined in order to fit the normalisation of the leading backgrounds
* Signal regions
  - Split according to number of leptons, total lepton charge, and a **multi-output deep neural network classifier** (DNN^{CAT}) to categorise the different BSM signals
  - A second **DNN** is trained in a distance correlation (disCo) approach in each CAT to discriminate the signal from the backgrounds (DNN^{SB})
* Simultaneous profile likelihood fit of CRs+SRs



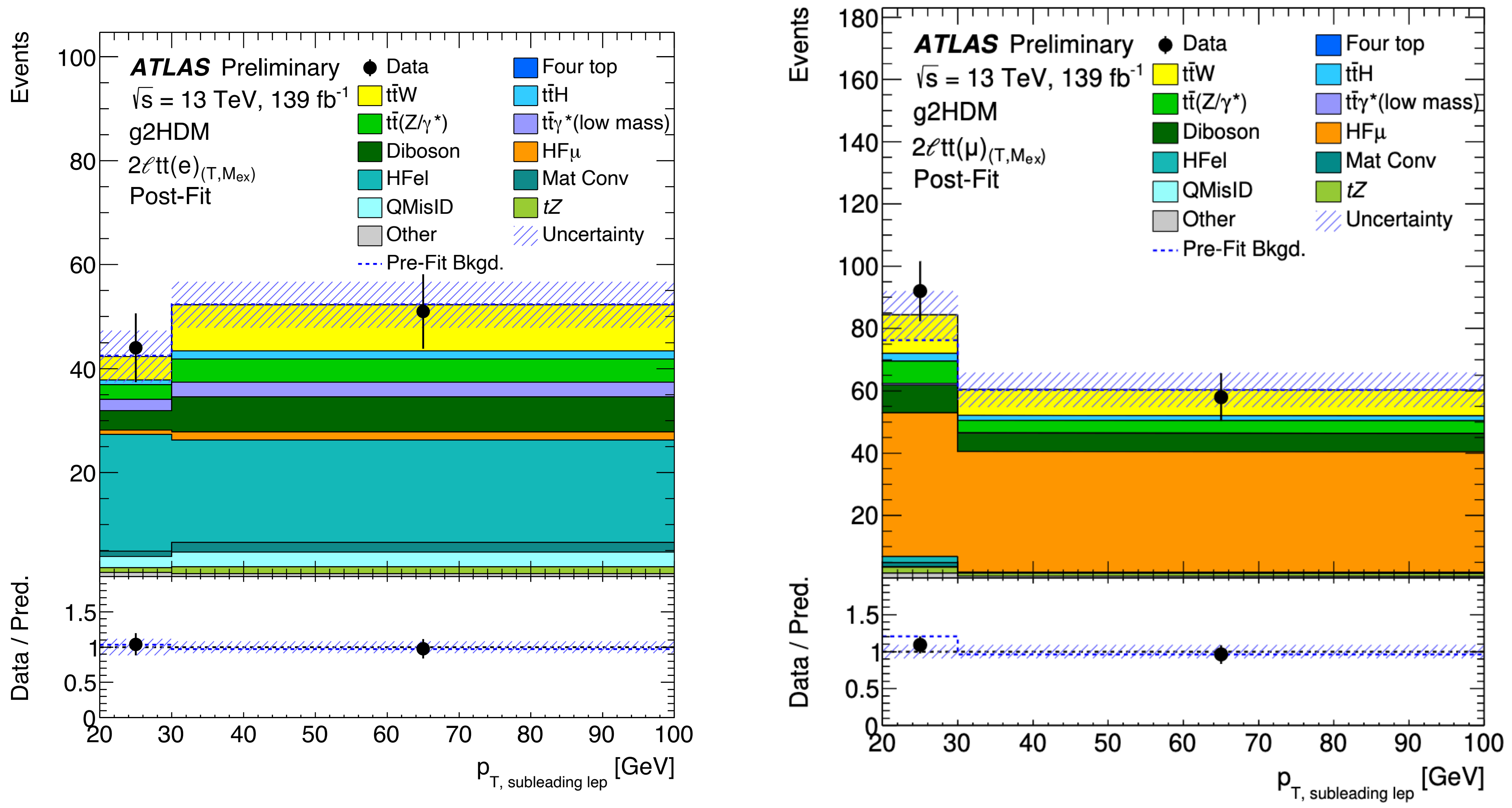
Analysis strategy

- Use **tight lepton definition** to suppress reducible background
  - Non-prompt lepton veto BDTs

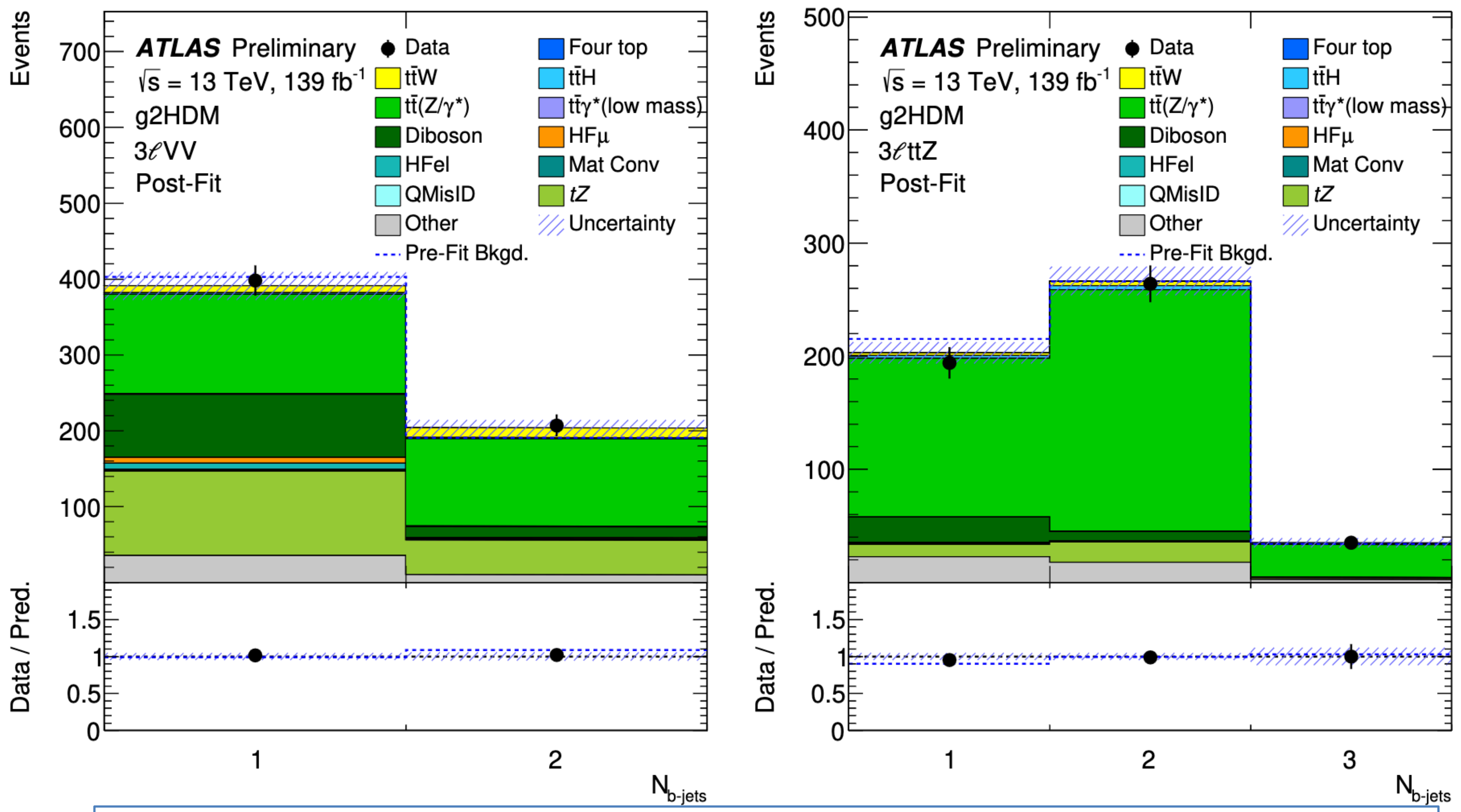
- **Estimate corrections** from data for various non-prompt ℓ (HF, conversions) and irreducible (VV+HF jets, $t\bar{t}Z$, $t\bar{t}W$) backgrounds in simultaneous fit CRs+SRs
  - Multiple **control regions** are defined in order to fit the normalisation of the leading backgrounds

- Signal regions
  - Split according to number of leptons, total lepton charge, and a **multi-output deep neural network classifier** (DNN^{CAT}) to categorise the different BSM signals
  - A second **DNN** is trained in a distance correlation (disCo) approach in each CAT to discriminate the signal from the backgrounds (DNN^{SB})

- Simultaneous profile likelihood fit of CRs+SRs



Data/MC comparison from CR dedicated for non-prompt electrons/muons from heavy-flavour decays



Good modeling of various Standard Model processes!

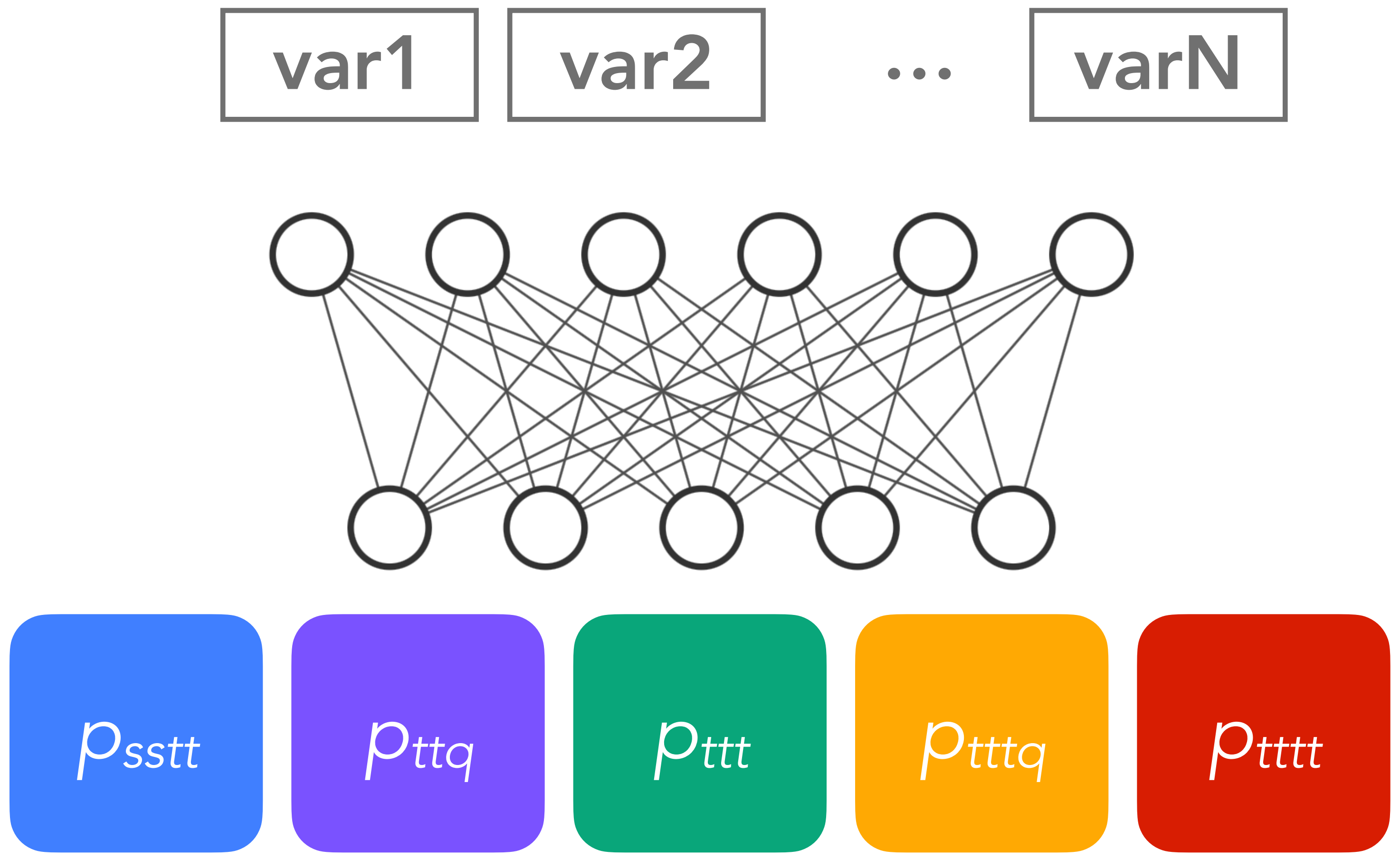
Data/MC comparison from CR dedicated for diboson and $t\bar{t}Z$ SM processes

Signal regions: Neural network(s)

1) (DNN^{CAT}) used to categorise the signal to various CATs has been trained separately for 2ℓSS & 3ℓ channels

- **2ℓSS**: 5 cat. (sstt, ttq, ttt, tttq, tttt)
- **3ℓ**: 3 cat. (ttt, tttq, tttt)

1st step



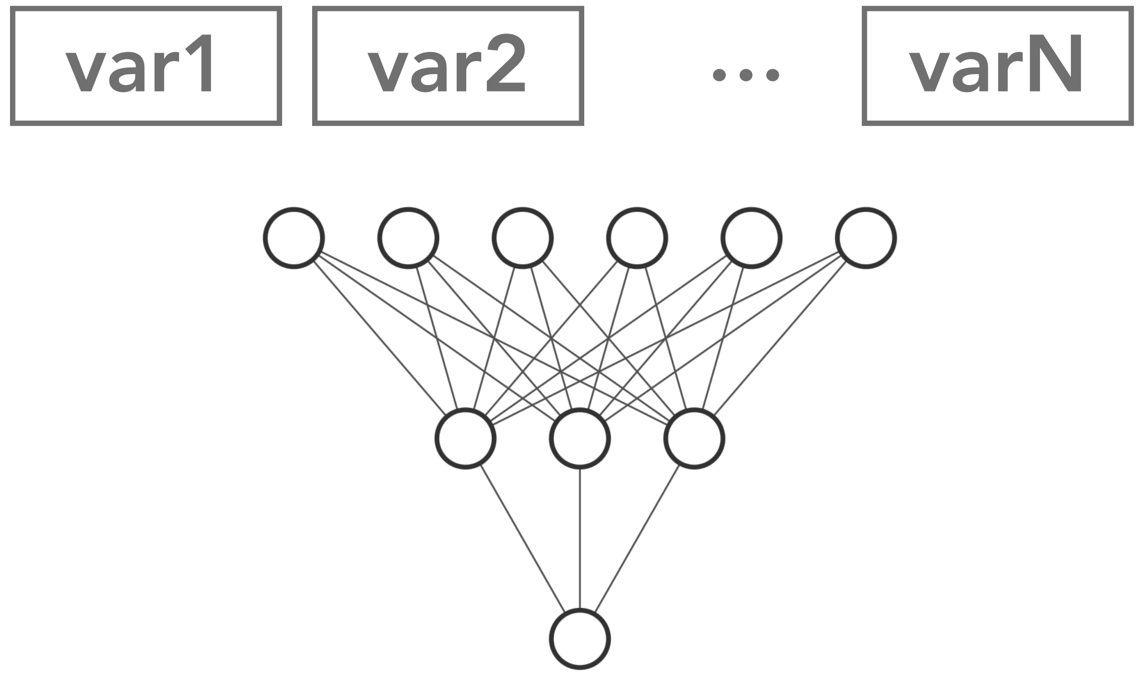
Class probabilities

2) (DNN^{SB}) used to discriminate the dedicated signal category from the background

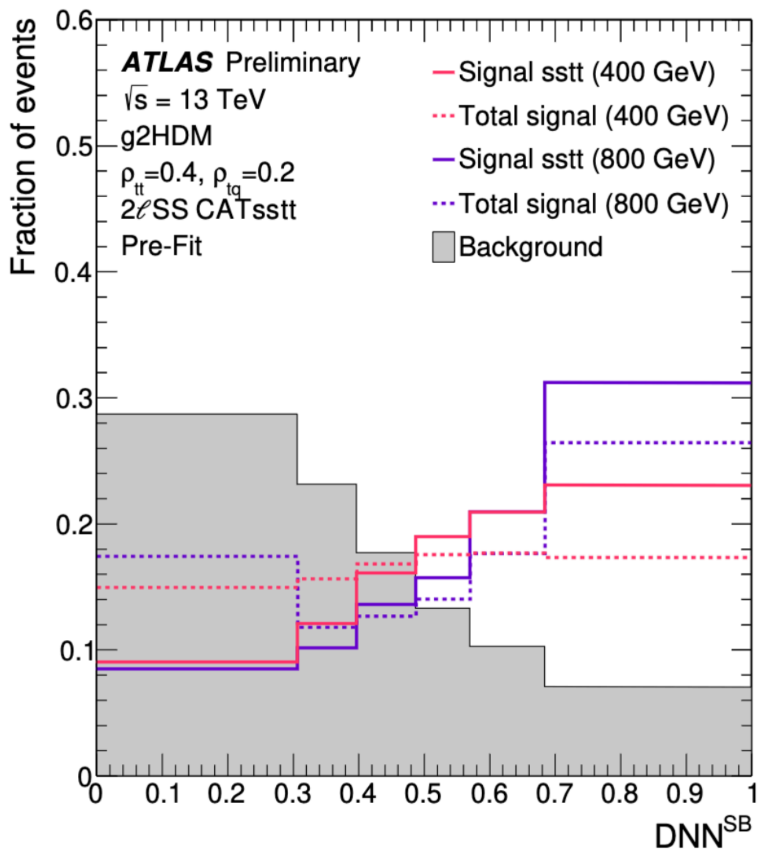
- **decorrelated** network from mass dependent variables (DisCo: Robust Networks Through Distance Correlation (arXiv))
- **2ℓSS**: 5 training, **3ℓ**: 3 training, **4ℓ**: 1 training

$$\ell = \ell_{\text{classifier}}(\vec{y}, \vec{y}_{true}) + \lambda \, \text{dCorr}^2(\vec{m}, \vec{y})$$

2nd step



CATs additionally split by charge to ++ and -- exploiting the additional separation between charge-symmetric and charge-asymmetric signal processes

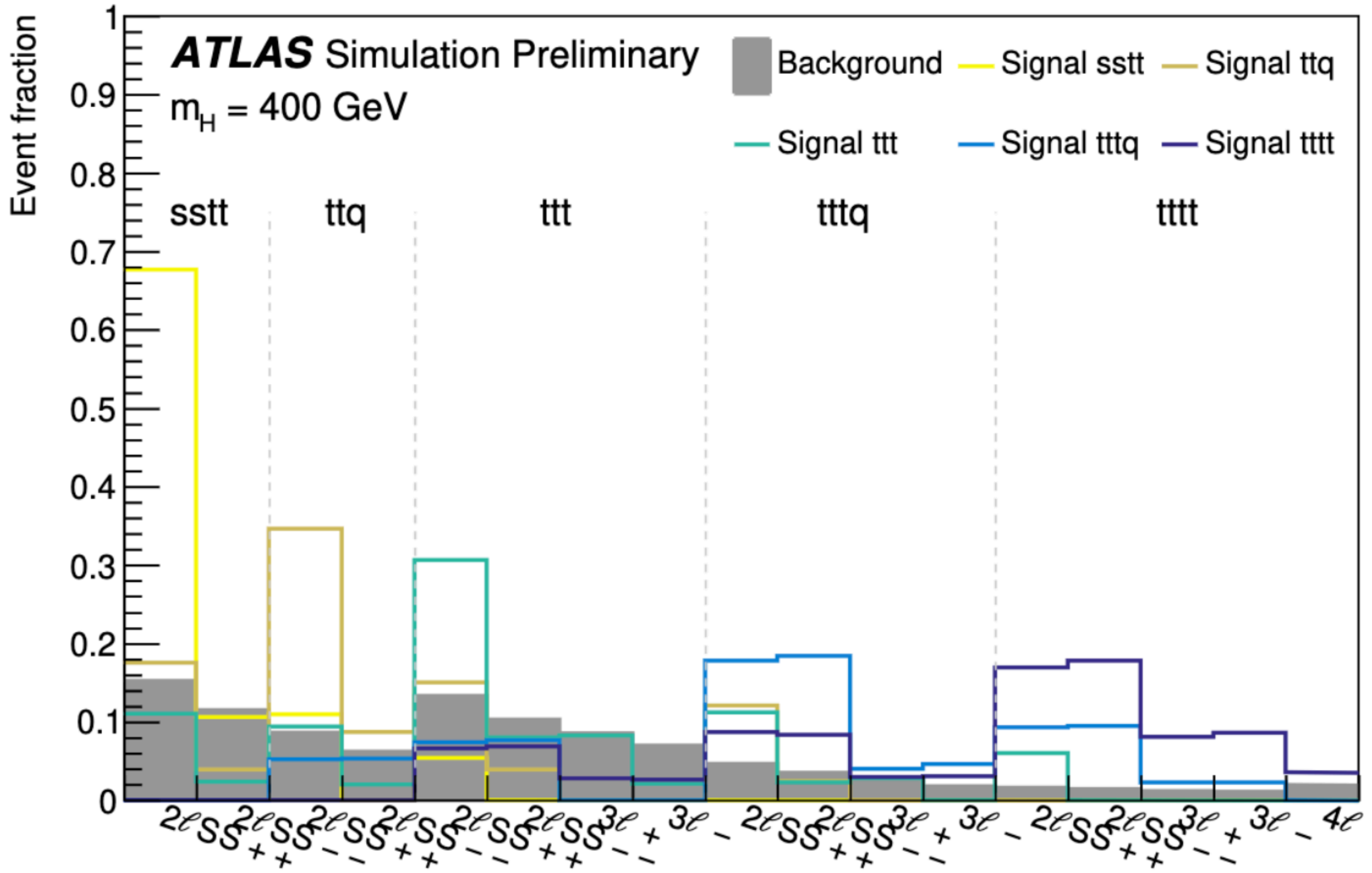


Variable	DNN^{cat}	DNN^{SB}
Number of jets (N_{jets})	✓	✓
Sum of pseudo-continuous b-tagging scores of jets	✓	✓
Pseudo-continuous b-tagging score of 1st, 2nd, 3rd leading jet in p_T	✓	✓
Sum of p_T of the jets and leptons ($H_{T,jets}$, $H_{T,lep}$)	✓	✓
Angular distance of leptons (sum in the case of 3ℓ and 4ℓ)	✓	✓
Missing transverse energy	✓	✓
Leading transverse momentum of jet	-	✓
Invariant mass of leading lepton and missing transverse energy	-	✓
Di/tri/quad-lepton type variable (associated to the number of electrons/muons in event)	-	✓

Signal regions: Neural network(s)

1) (DNN^{CAT}) used to categorise the signal to various CATs has been trained separately for 2ℓSS & 3ℓ channels

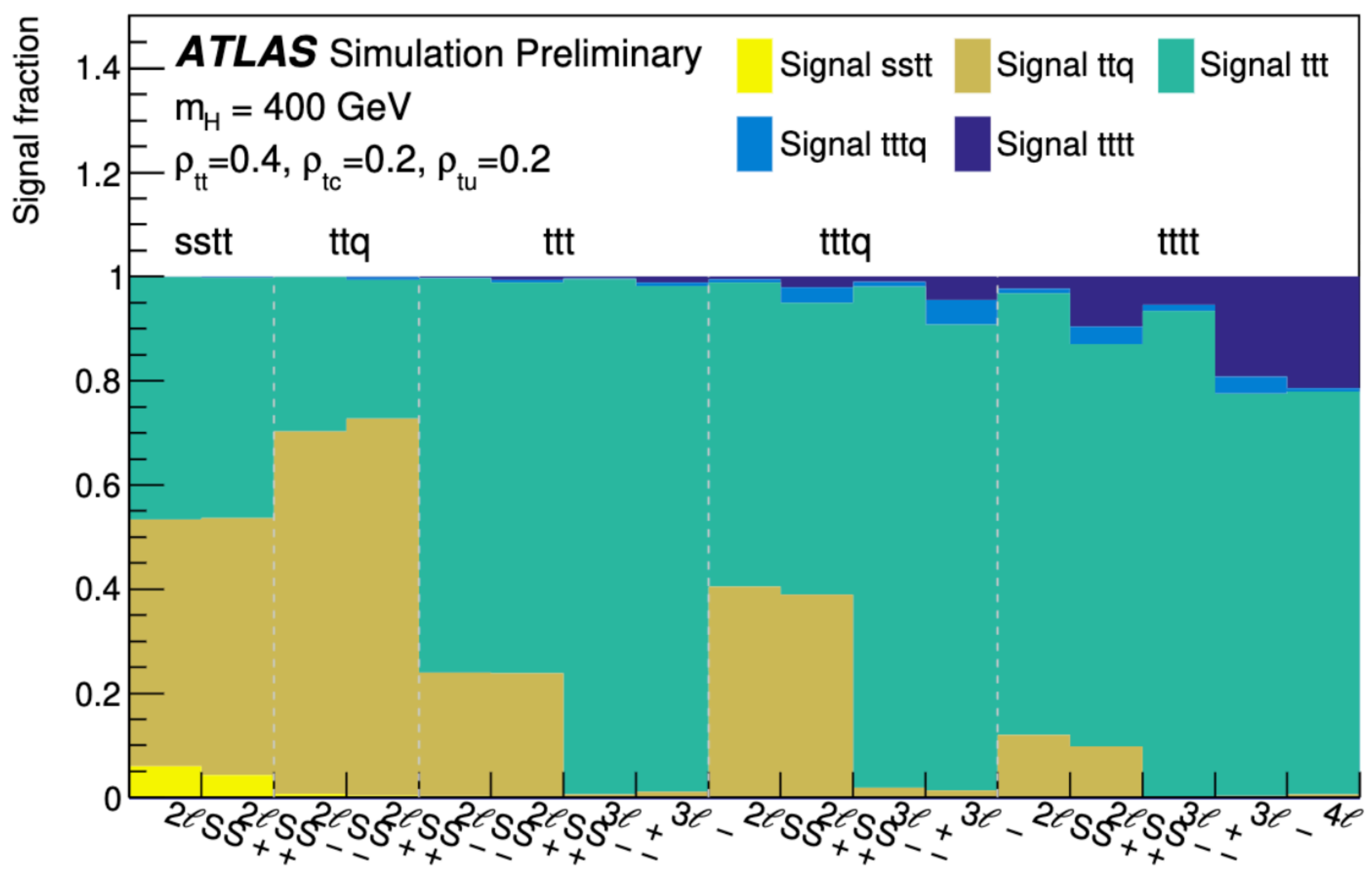
- **2ℓSS**: 5 cat. (sstt, ttq, ttt, tttq, tttt)
- **3ℓ**: 3 cat. (ttt, tttq, tttt)



Normalized distributions of the categorization for each signal process, compared to the total expected background,

2) (DNN^{SB}) used to discriminate the dedicated signal category from the background

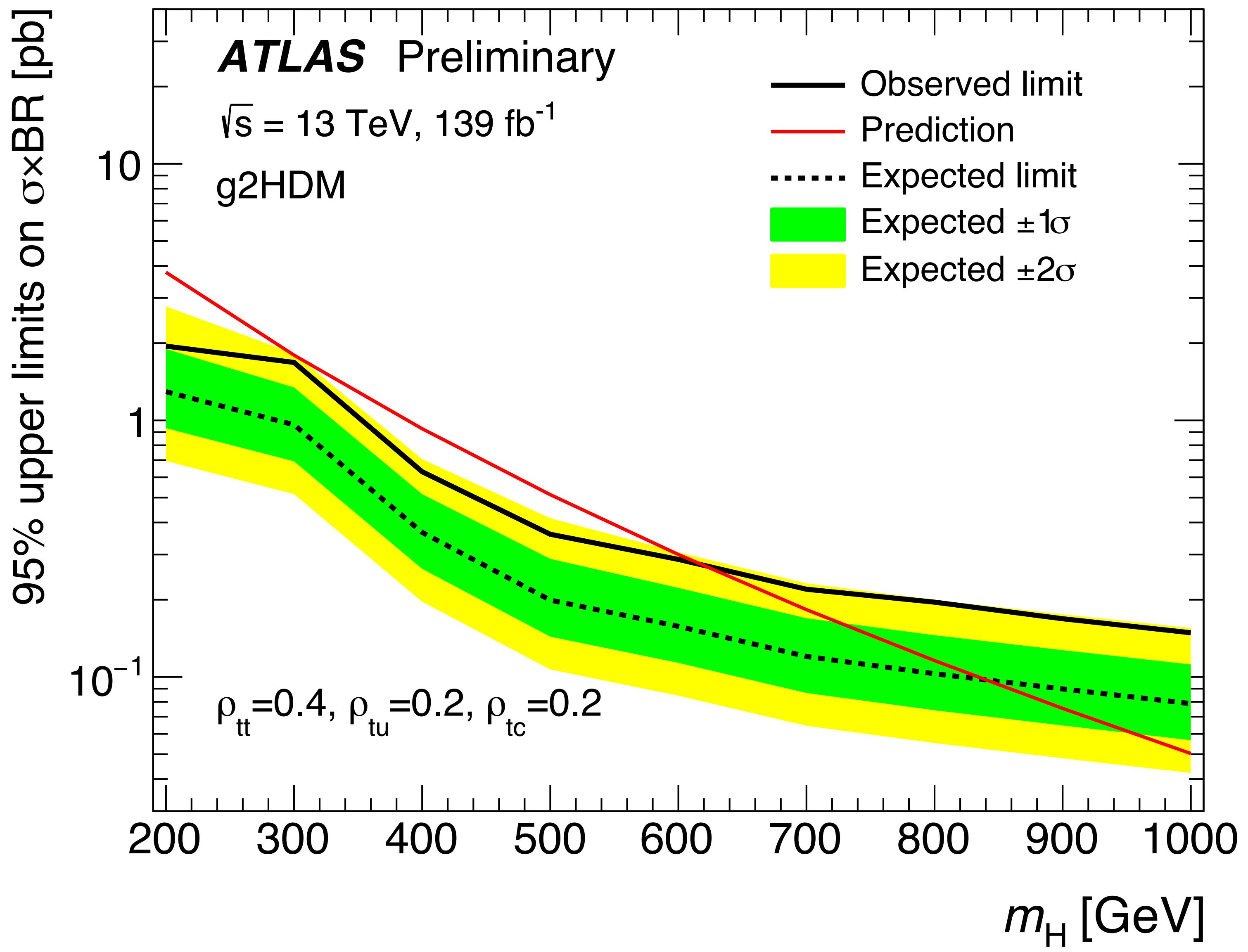
- **decorrelated** network from mass dependent variables (DisCo: Robust Networks Through Distance Correlation (arXiv))
- **2ℓSS**: 5 training, **3ℓ**: 3 training, **4ℓ**: 1 training



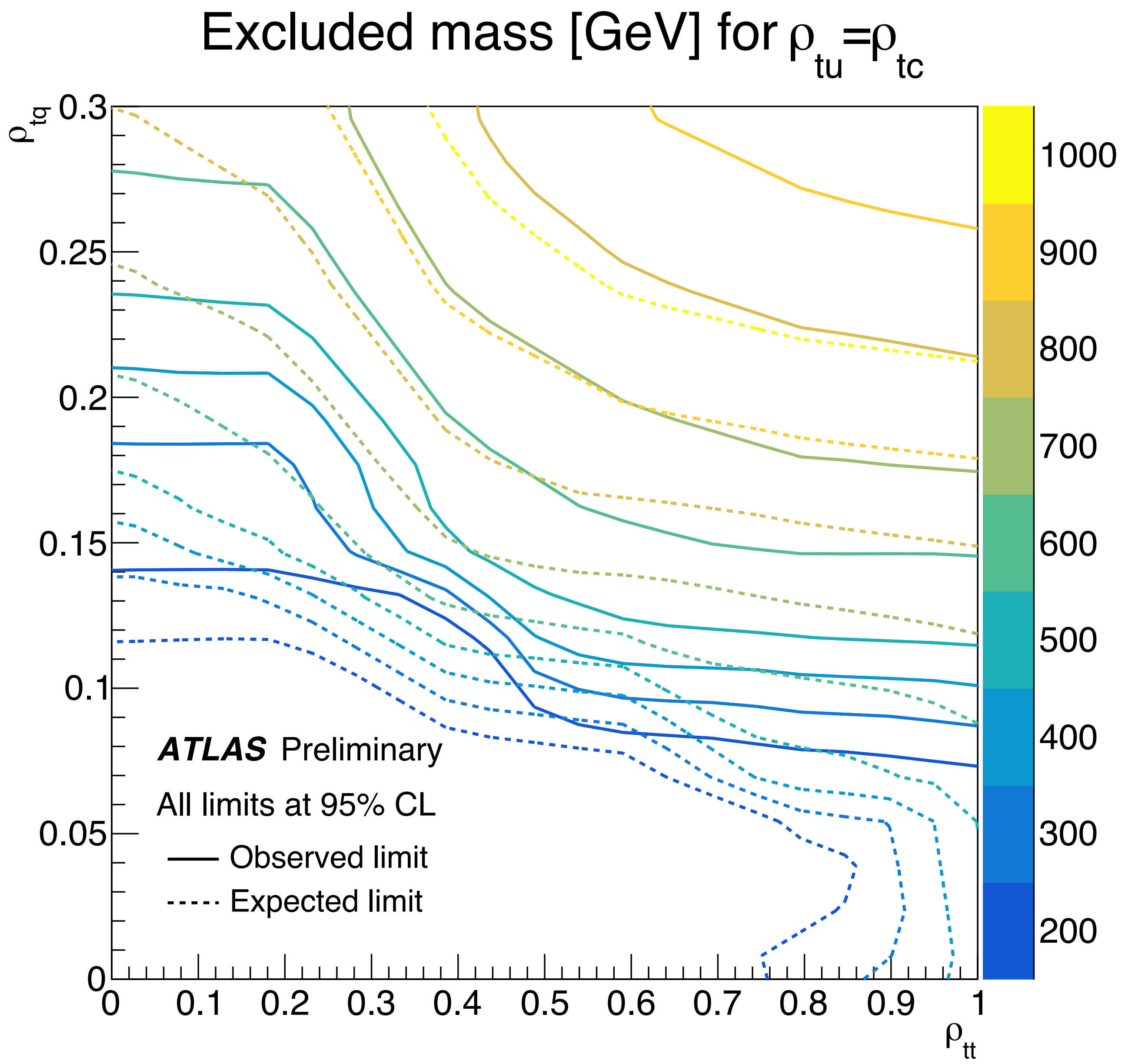
The expected fractional signal contribution in each category scalar masses for the coupling set ρ_{tt} = 0.4, ρ_{tc} = 0.2, ρ_{tu} = 0.2.

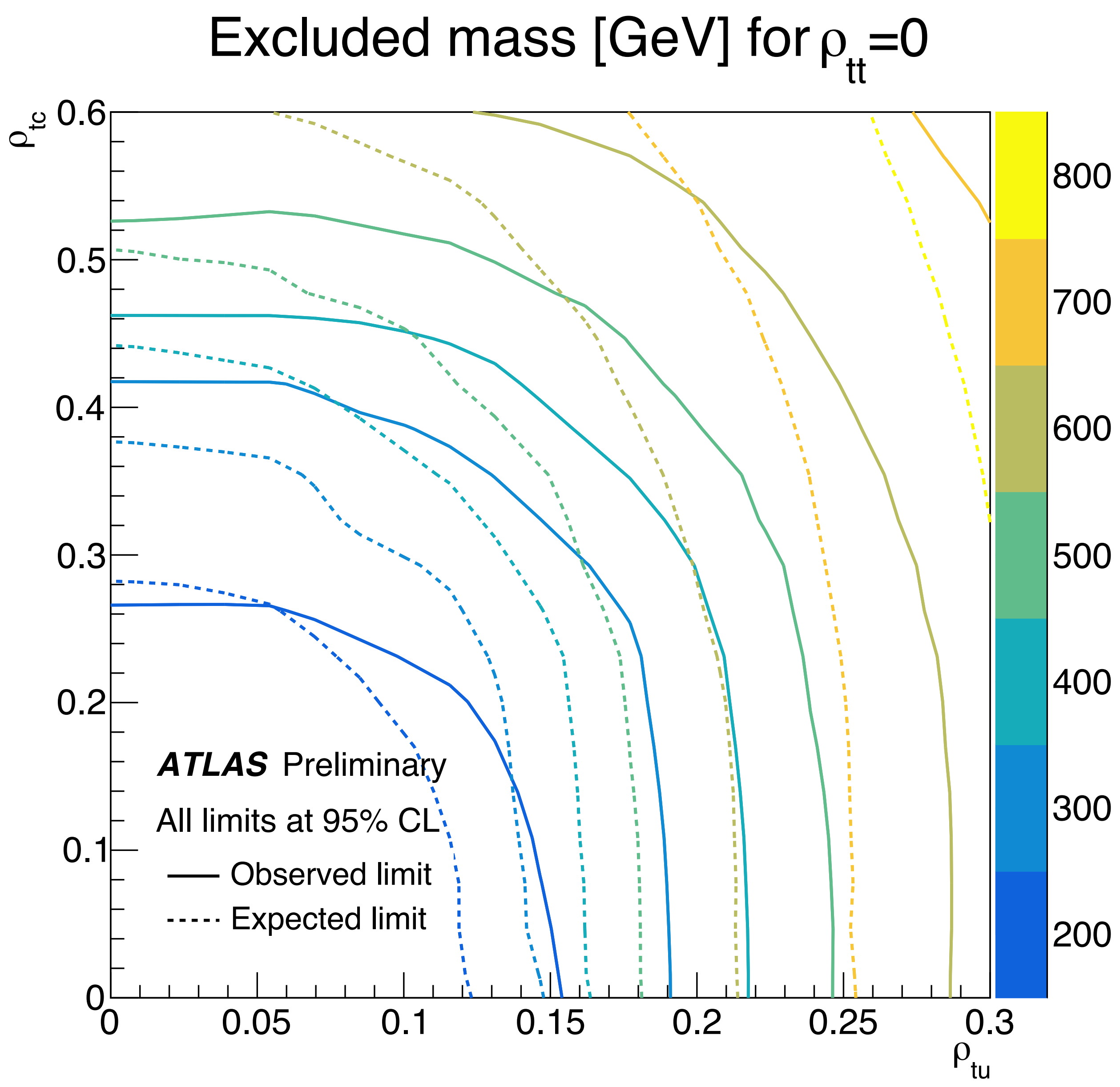
Results

* **95% CL upper limits** on the **σ x BR** for all the 2HDM signals together across different heavy Higgs masses (from 200 to 1000 GeV)
    - Assuming BSM signal corresponding to benchmark match $t\bar{t}W$ and 4-tops tensions: 400 GeV and $\rho_{tt} = 0.4,\ \rho_{tc} = 0.2,\ \rho_{tu} = 0.2$
    - Assuming $\rho_{tt} = 0$ **or** $\rho_{tq} = 0$



Observed exclusion limits on the σ x BR ratio for the benchmark coupling scenario $m_H < 630\ GeV$ excluded

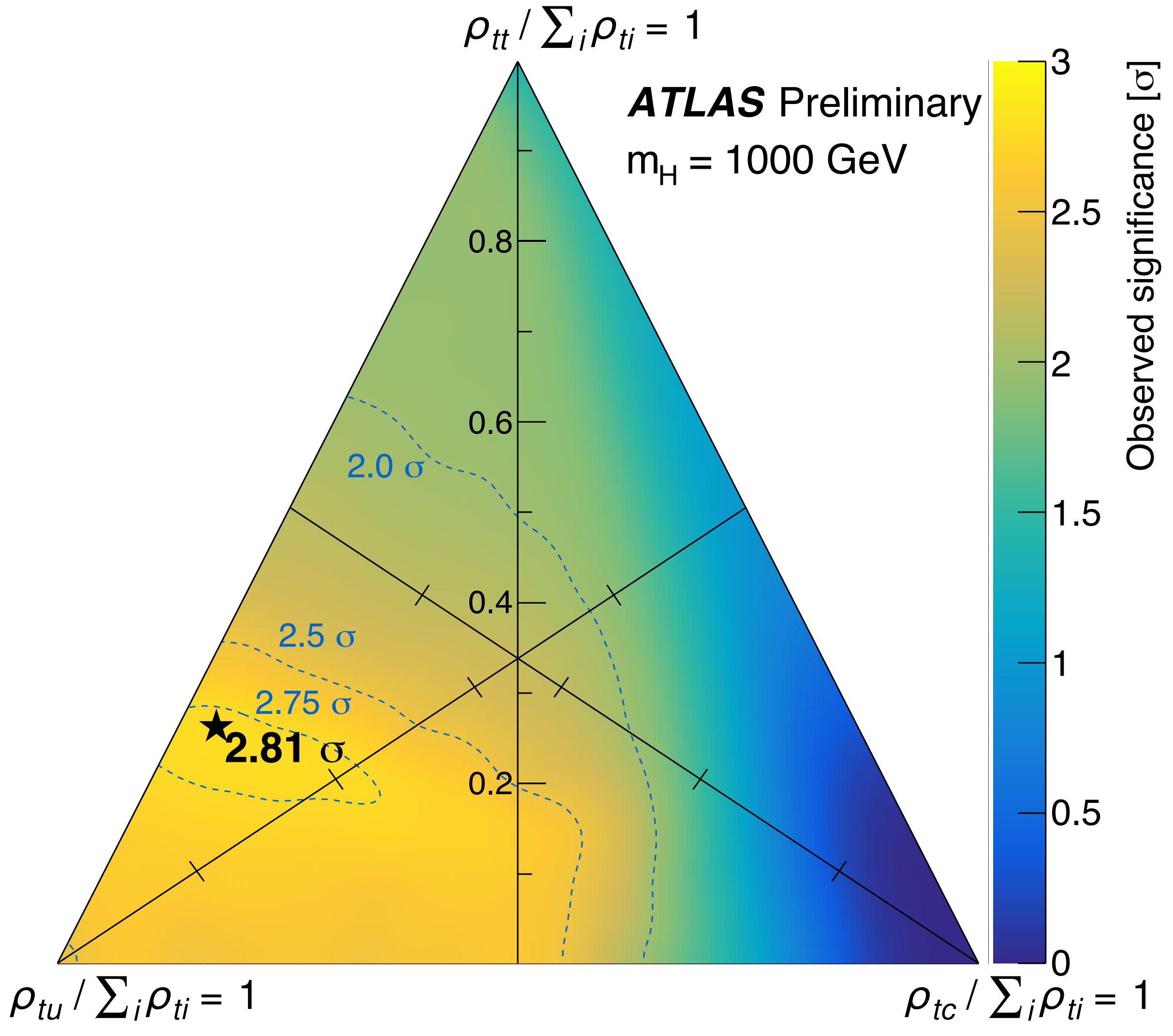




Observed (solid line) and expected (dashed line) exclusion limits on the scalar mass as a function of the coupling under $\rho_{tc} = \rho_{tu}$ (left) and $\rho_{tt} = 0$ (right) scenarios

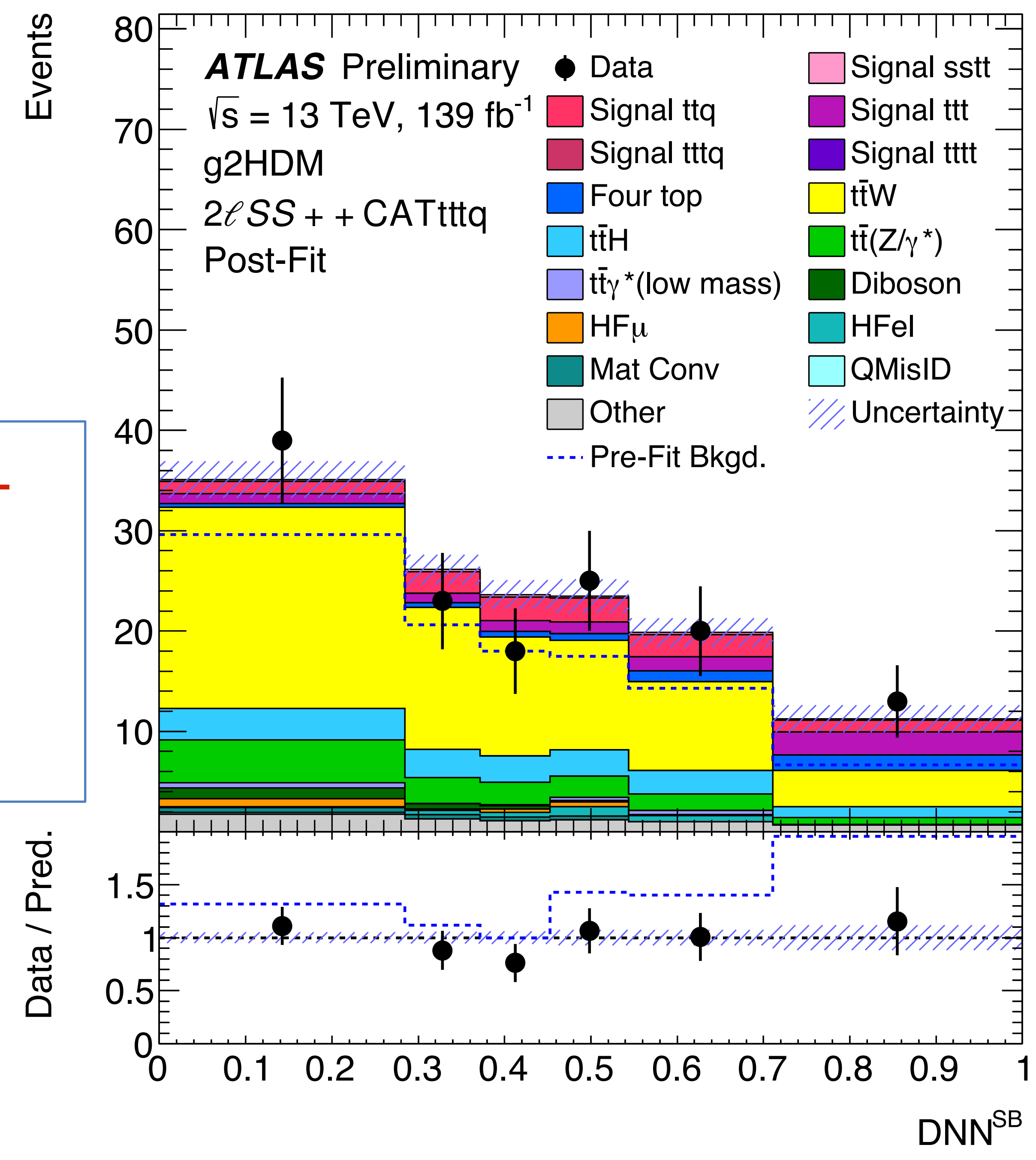
Results

- 95% CL upper limits on the σ x BR for all the 2HDM signals together across different heavy Higgs masses (from 200 to 1500 GeV)
  - Assuming BSM signal corresponding to benchmark match $t\bar{t}W$ and 4-tops tensions: 400 GeV and $\rho_{tt} = 0.4,\ \rho_{tc} = 0.2,\ \rho_{tu} = 0.2$
  - Assuming $\rho_{tt} = 0$ **or** $\rho_{tq} = 0$
- **Scan the full 4D planes of couplings** $(\rho_{tt},\ \rho_{tu}\ ,\rho_{tc})$ **vs mass**
  - **A mild excess is observed** over SM expectation to a local significance of **2.81**σ for a signal with $m_H = 1000\ GeV$ and $(\rho_{tt} = 0.32,\ \rho_{tc} = 0.05,\ \rho_{tu} = 0.85)$

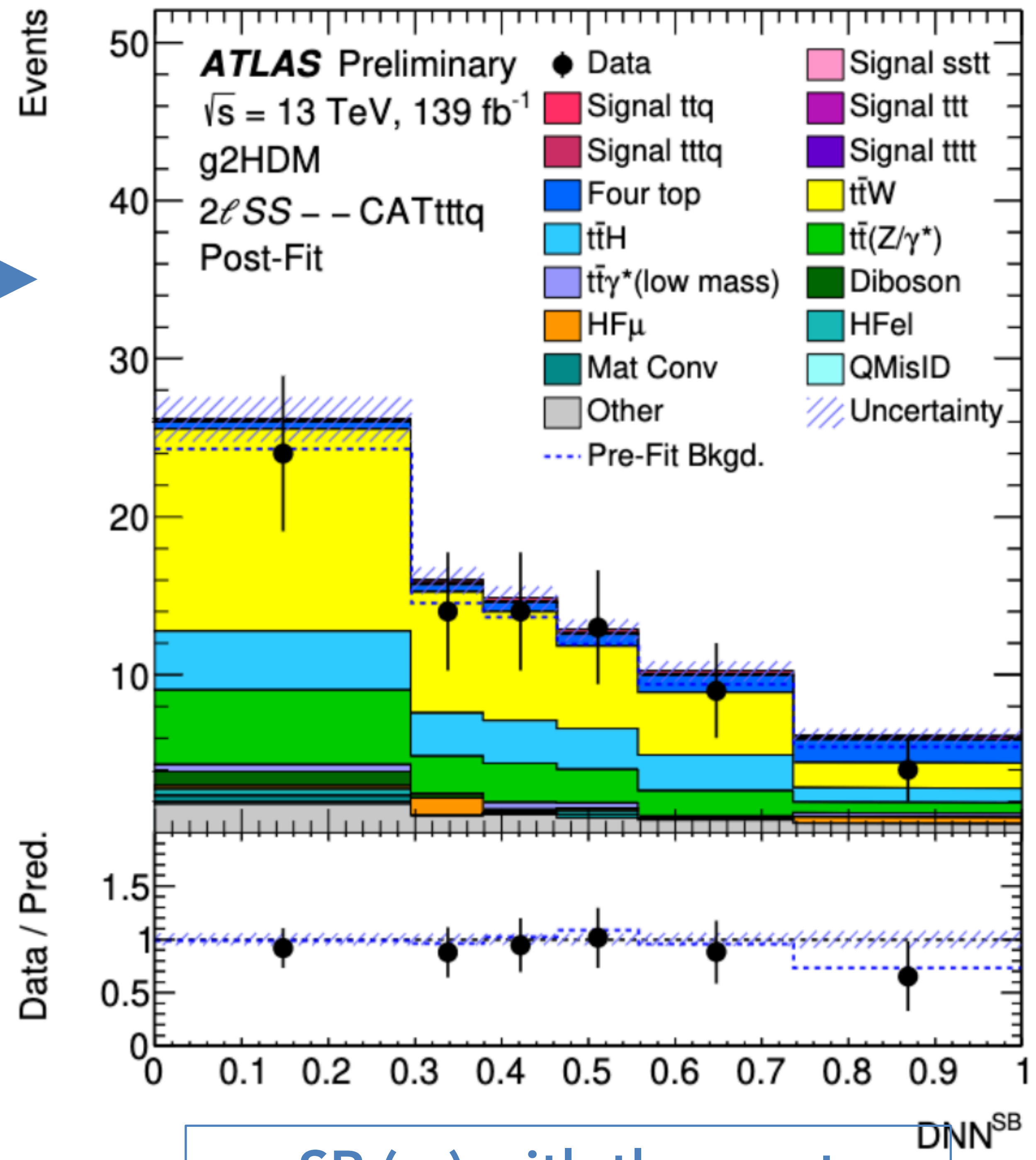


Scan of the couplings at $m_H = 1\ TeV$

Excess in ++ ttq/tttq categories



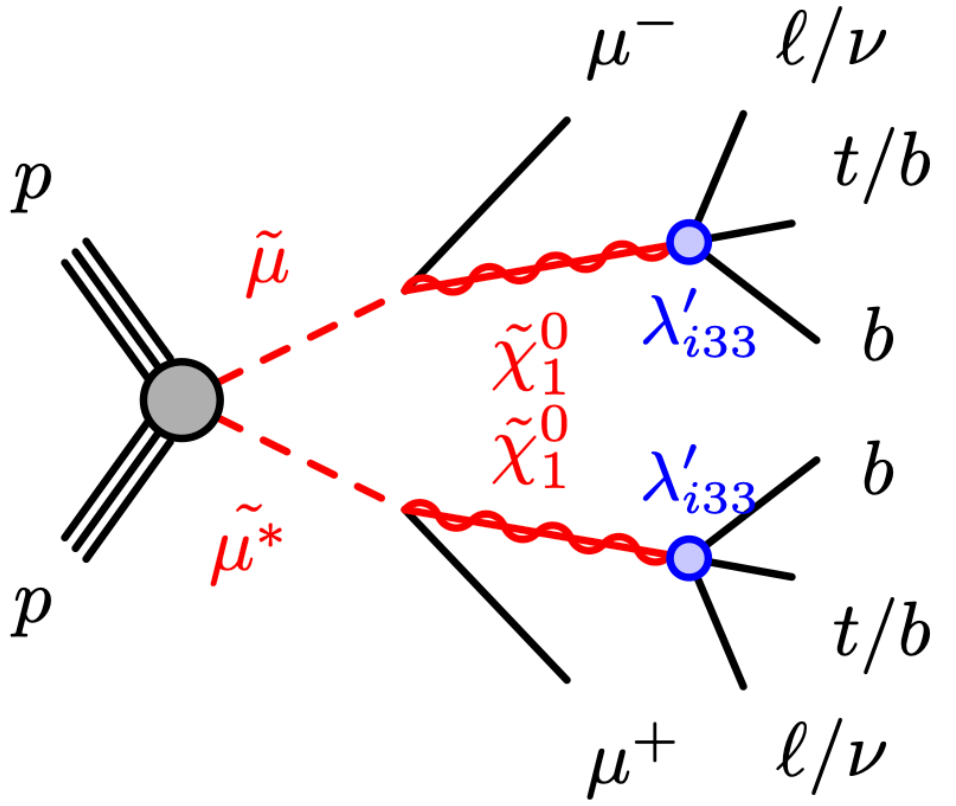
SR (++) with the most significant tension

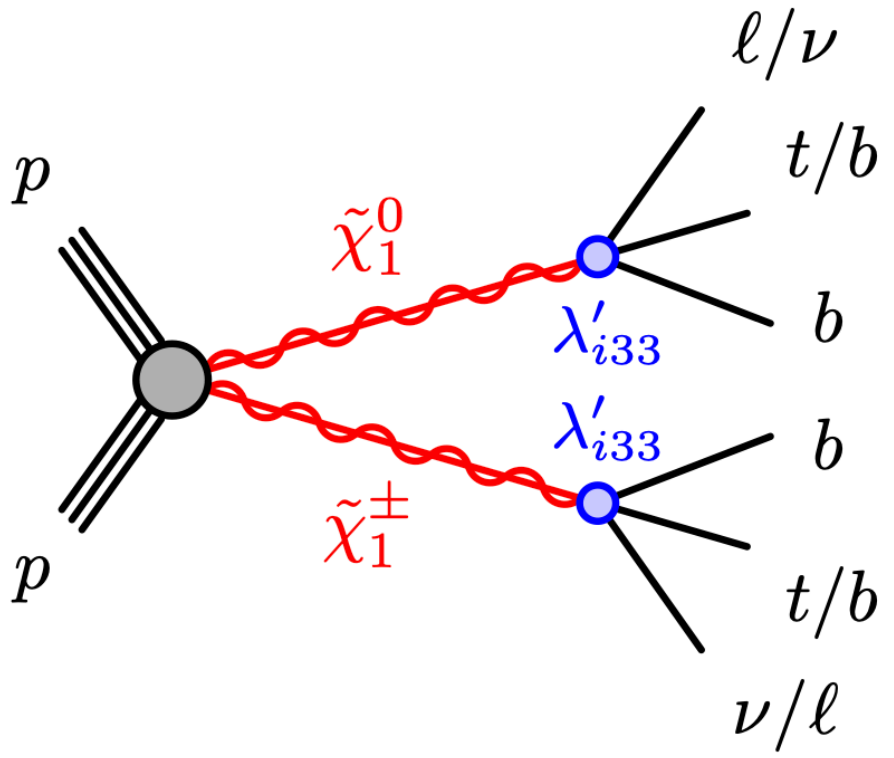


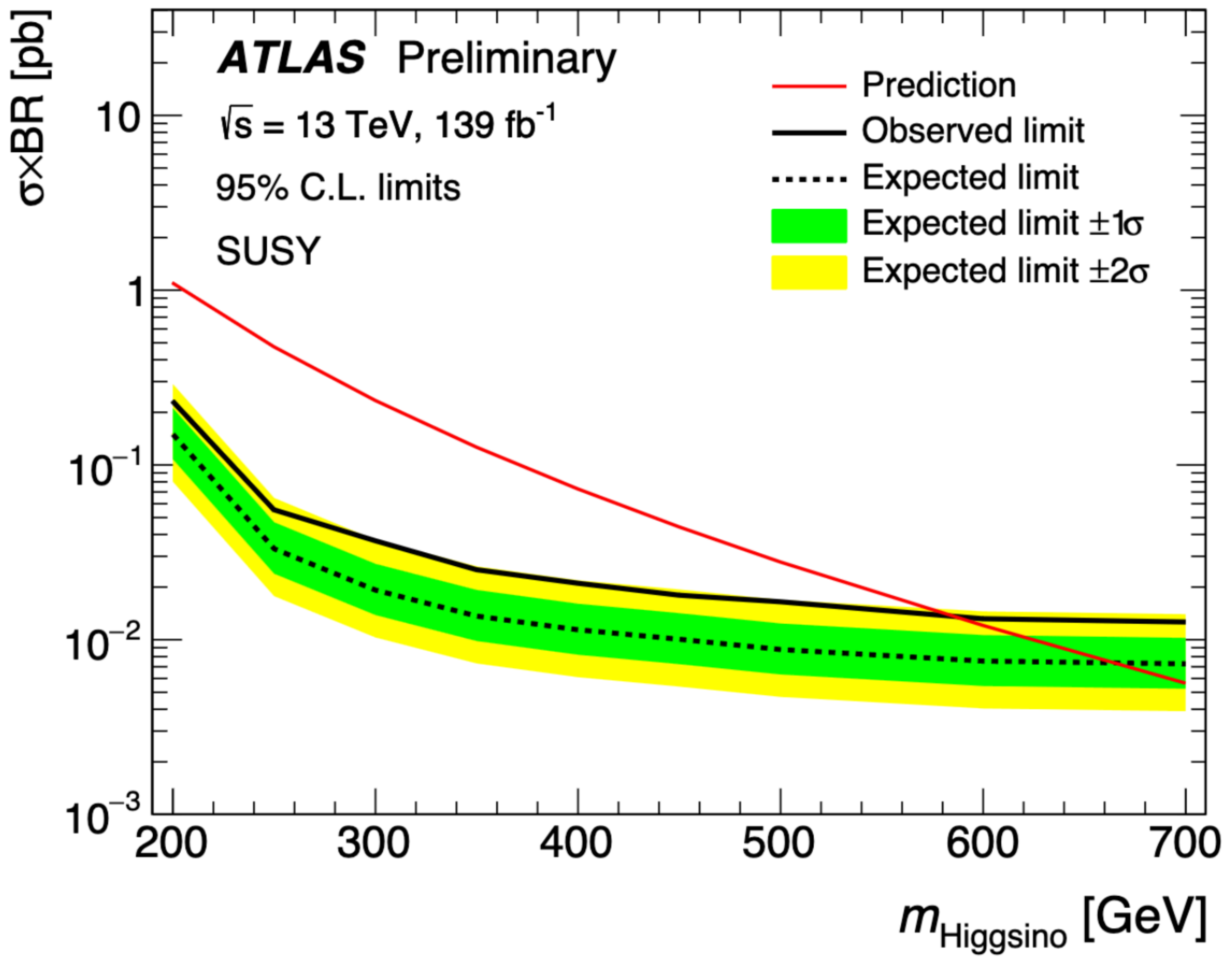
SR (- -) with the most significant tension

Results

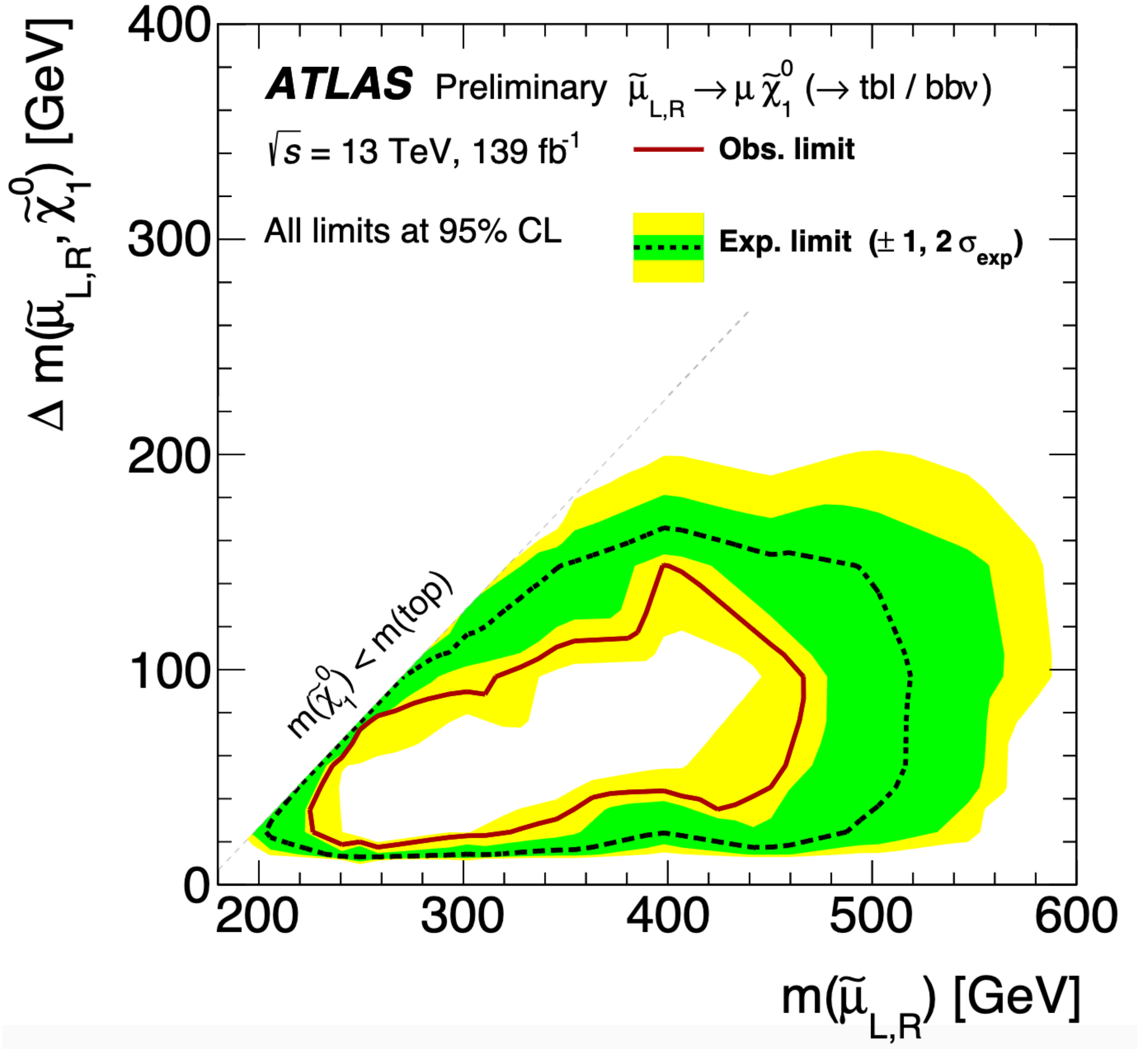
- 95% CL upper limits on the σ x BR for all the 2HDM signals together across different heavy Higgs masses (from 200 to 1500 GeV)
  - Assuming BSM signal corresponding to benchmark match $t\bar{t}W$ and 4-tops tensions: 400 GeV and $\rho_{tt} = 0.4,\ \rho_{tc} = 0.2,\ \rho_{tu} = 0.2$
  - Assuming $\rho_{tt} = 0$ **or** $\rho_{tq} = 0$
- **Scan the full 4D planes of couplings** $(\rho_{tt},\ \rho_{tu}\ ,\rho_{tc})$ **vs mass**
  - **A mild excess is observed** over SM expectation to a local significance of **2.81**σ for a signal with $m_H = 1000\ GeV$ and $(\rho_{tt} = 0.32,\ \rho_{tc} = 0.05,\ \rho_{tu} = 0.85)$
- **Bonus:** RPV SUSY interpretation (decay via a lepton-number- violating RPV coupling)
  - Set limits on wino, higgsino and smuon-bino models with RPV LQD coupling







Observed exclusion limits on the σ x BR ratio for Higgsino model $200 < m_H < 585\ GeV$ excluded



Observed exclusion limits on the σ x BR ratio for Smuon-bino model $225 < m_H < 600\ GeV$ excluded

Summary

* Analysis search for heavy Higgs bosons from a g2HDM in multilepton plus b-jets final states in pp collisions at 13 TeV with the ATLAS detector
  * Presented in conference note - Link
* Additional sensitivity to novel SUSY signals
* A mild excess is observed over the Standard Model expectation corresponding to a local significance of 2.81 standard deviations for a signal with $m_H = 1000\ GeV$ and $(\rho_{tt} = 0.32,\ \rho_{tc} = 0.05,\ \rho_{tu} = 0.85)$
* BSM signal profiting mostly from charge-asymmetric ++ ttq/tttq categories
* **New paper with updated mass points** is on the way, stay tuned!

Backup

Backup

Table 2: Description of the loose inclusive ("L"), medium inclusive ("M"), medium exclusive ("M_{ex}"), and tight ("T") lepton definitions. The electron e^* is required to fulfil, in addition to the corresponding lepton definition requirements, those corresponding to an internal or material conversion candidate.

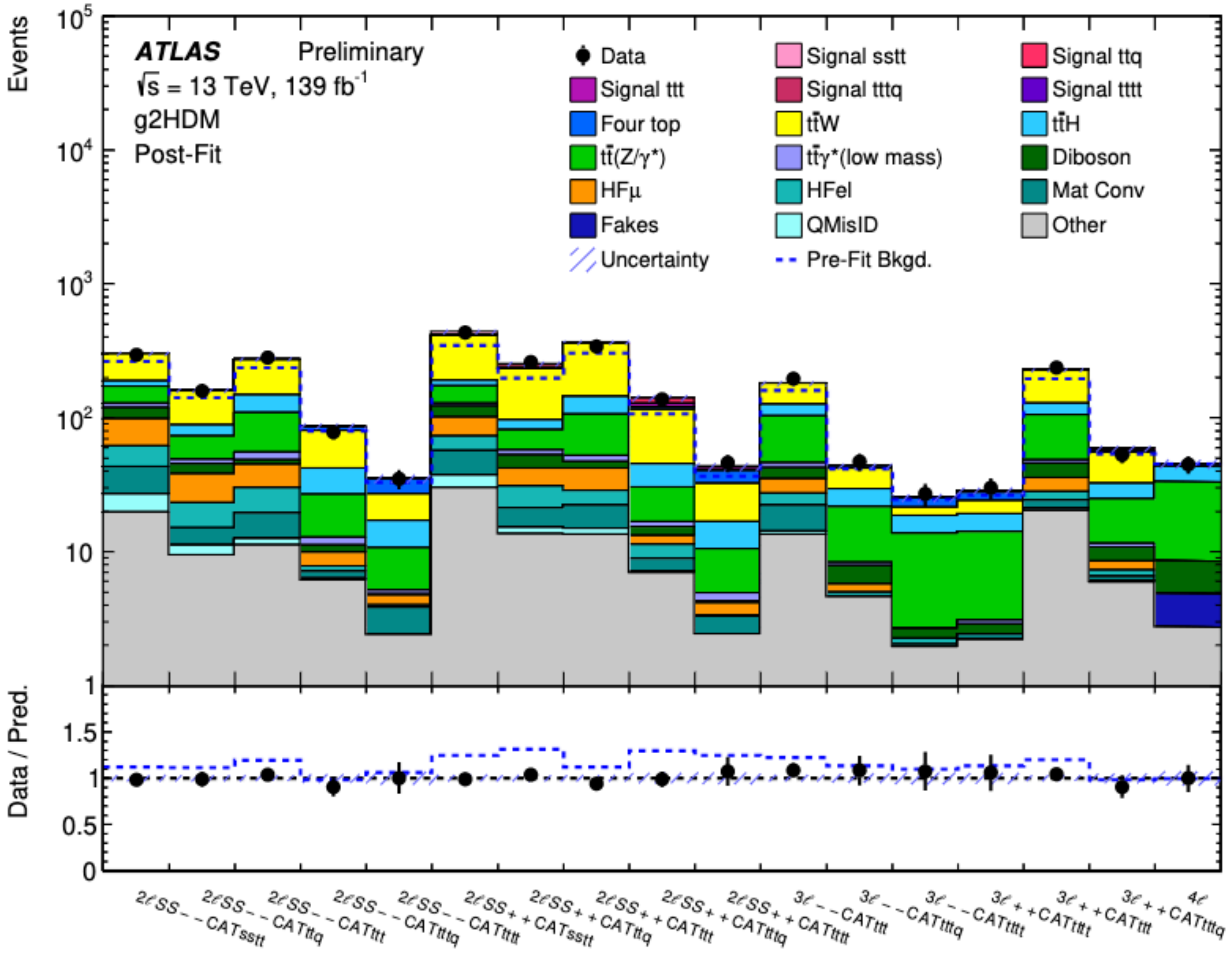
	e				μ			
	L	M	M_{ex}	T	L	M	M_{ex}	T
Isolation	Yes				Yes			
Non-prompt lepton BDT WP	No	*Tight*	*Tight-not-VeryTight*	*VeryTight*	No	*Tight*	*Tight-not-VeryTight*	*VeryTight*
Identification	Loose	Tight			Loose	Medium		
Electron charge-misassignment veto	No	Yes			N/A			
Electron conversion candidate veto	No	Yes (except e^*)			N/A			
Transverse impact parameter significance $\lvert d_0 \rvert / \sigma_{d_0}$	< 5				< 3			
Longitudinal impact parameter $\lvert z_0 \sin\theta \rvert$	< 0.5 mm							

Backup

Table 4: Event selection summary in the control regions. The notation e^* is used to denote material conversion or internal conversion candidates, as described in Section 4. In the HF non-prompt lepton region naming, “2ℓSStt(e)” (“2ℓSStt(μ)”) refer to the control regions enriched in non-prompt electrons (muons) from semileptonic b-decays originating mostly from $t\bar{t}$ and with the lepton flavours for the leading and subleading leptons corresponding to “ee, μe” (“$\mu\mu$, $e\mu$”). The additional (T, M_{ex}), (M_{ex}, T), and $(M_{\text{ex}}, M_{\text{ex}})$ subscripts refer to the lepton definitions required for the leading and subleading leptons in each region.

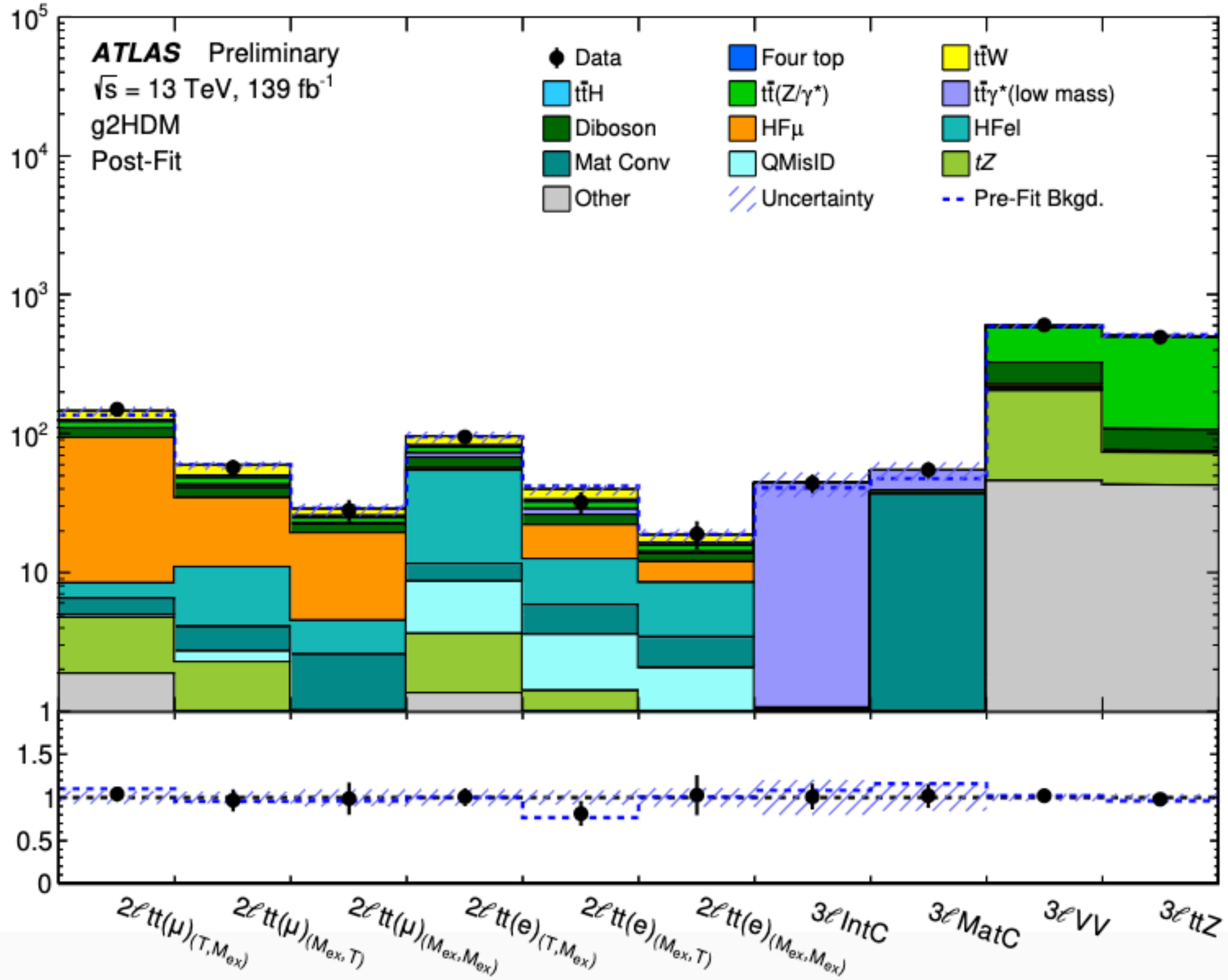
Control regions	WZ	$t\bar{t}Z$	Conversions	HF non-prompt
N_{jets}	2 or 3	≥ 4	≥ 0	≥ 2
$N_{b-\text{jets}}$	$\geq 1\ b^{60\%} \parallel \geq 2\ b^{77\%}$		$0\ b^{77\%}$	$1\ b^{77\%}$
Lepton requirement	3ℓ		$\mu\mu e^*$	2ℓSS
Lepton definition	(L, M, M)			$(T, M_{\text{ex}}) \parallel (M_{\text{ex}}, T) \parallel (M_{\text{ex}}, M_{\text{ex}})$
Lepton p_{T} [GeV]	(10, 20, 20)			(20, 20)
$m_{\ell^+\ell^-}^{OS-SF}$ [GeV]	>12		>12	–
$\lvert m_{\ell^+\ell^-}^{OS-SF} - m_Z \rvert$ [GeV]	<10		>10	–
$\lvert m_{\ell\ell\ell} - m_Z \rvert$ [GeV]	–		<10	–
$m_T(\ell_0, E_{\text{T}}^{\text{miss}})$ [GeV]	–			< 250
Region split	–	–	internal / material	subleading $e/\mu \times [(T, M_{\text{ex}}), (M_{\text{ex}}, T), (M_{\text{ex}}, M_{\text{ex}})]$
Region naming	3ℓVV	3ℓttZ	3ℓIntC	$2\ell\text{tt(e)}_{(T,M_{\text{ex}})}$, $2\ell\text{tt(e)}_{(M_{\text{ex}},T)}$, $2\ell\text{tt(e)}_{(M_{\text{ex}},M_{\text{ex}})}$
			3ℓMatC	$2\ell\text{tt}(\mu)_{(T,M_{\text{ex}})}$, $2\ell\text{tt}(\mu)_{(M_{\text{ex}},T)}$, $2\ell\text{tt}(\mu)_{(M_{\text{ex}},M_{\text{ex}})}$

Backup



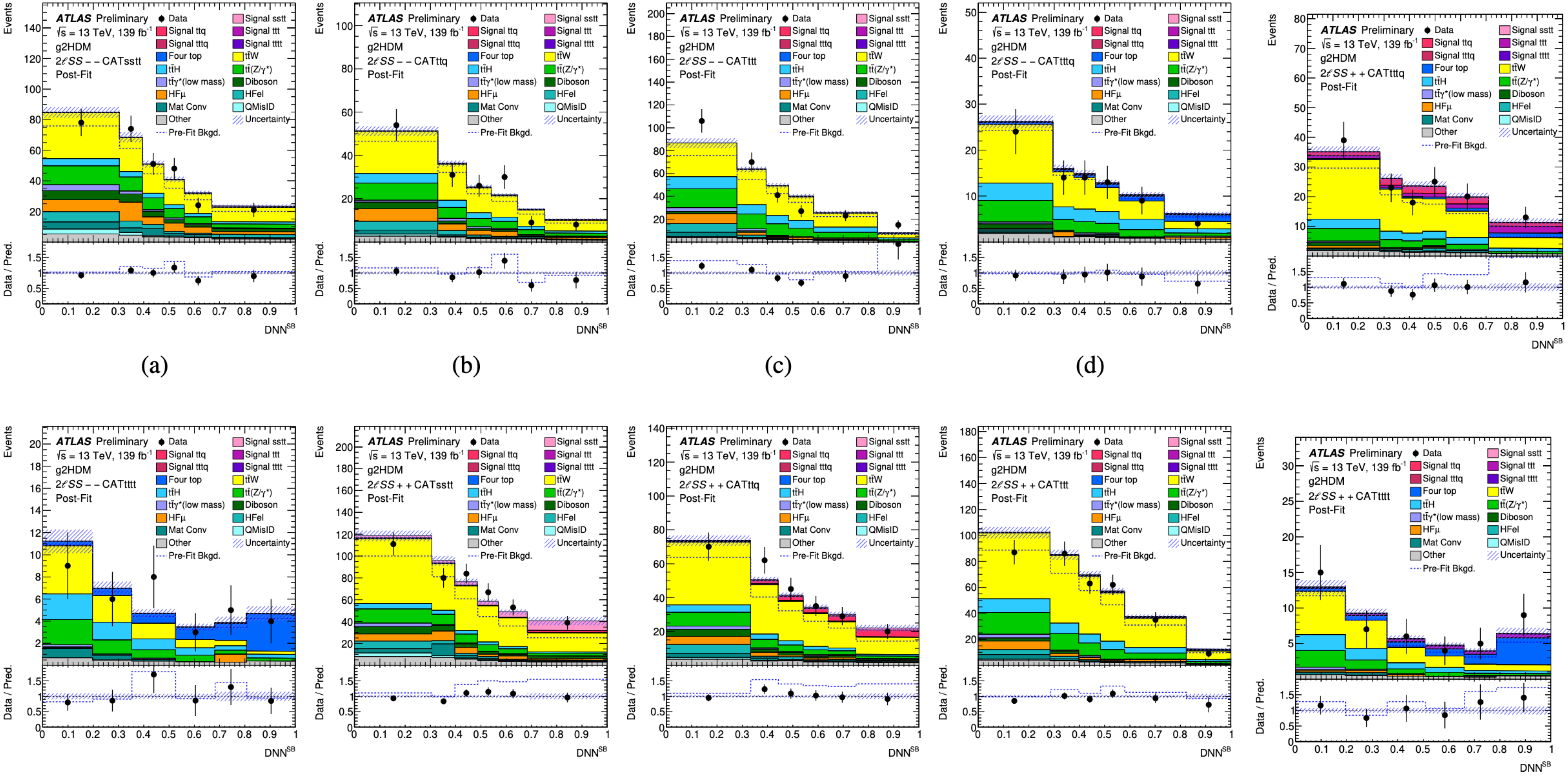
ATLAS Preliminary
√s = 13 TeV, 139 fb-1
g2HDM
Post-Fit
Data
Signal sstt
Signal ttq
Signal ttt
Signal tttq
Signal tttt
Four top
ttW
ttH
tt(Z/γ*)
ttγ*(low mass)
Diboson
HFμ
HFel
Mat Conv
Fakes
QMisID
Other
Uncertainty
Pre-Fit Bkgd.
Events
Data / Pred.

(a)



ATLAS Preliminary
√s = 13 TeV, 139 fb-1
g2HDM
Post-Fit
Data
Four top
ttW
ttH
tt(Z/γ*)
ttγ*(low mass)
Diboson
HFμ
HFel
Mat Conv
QMisID
tZ
Other
Uncertainty
Pre-Fit Bkgd.
3ℓ IntC
3ℓ MatC
3ℓ VV
3ℓ ttZ

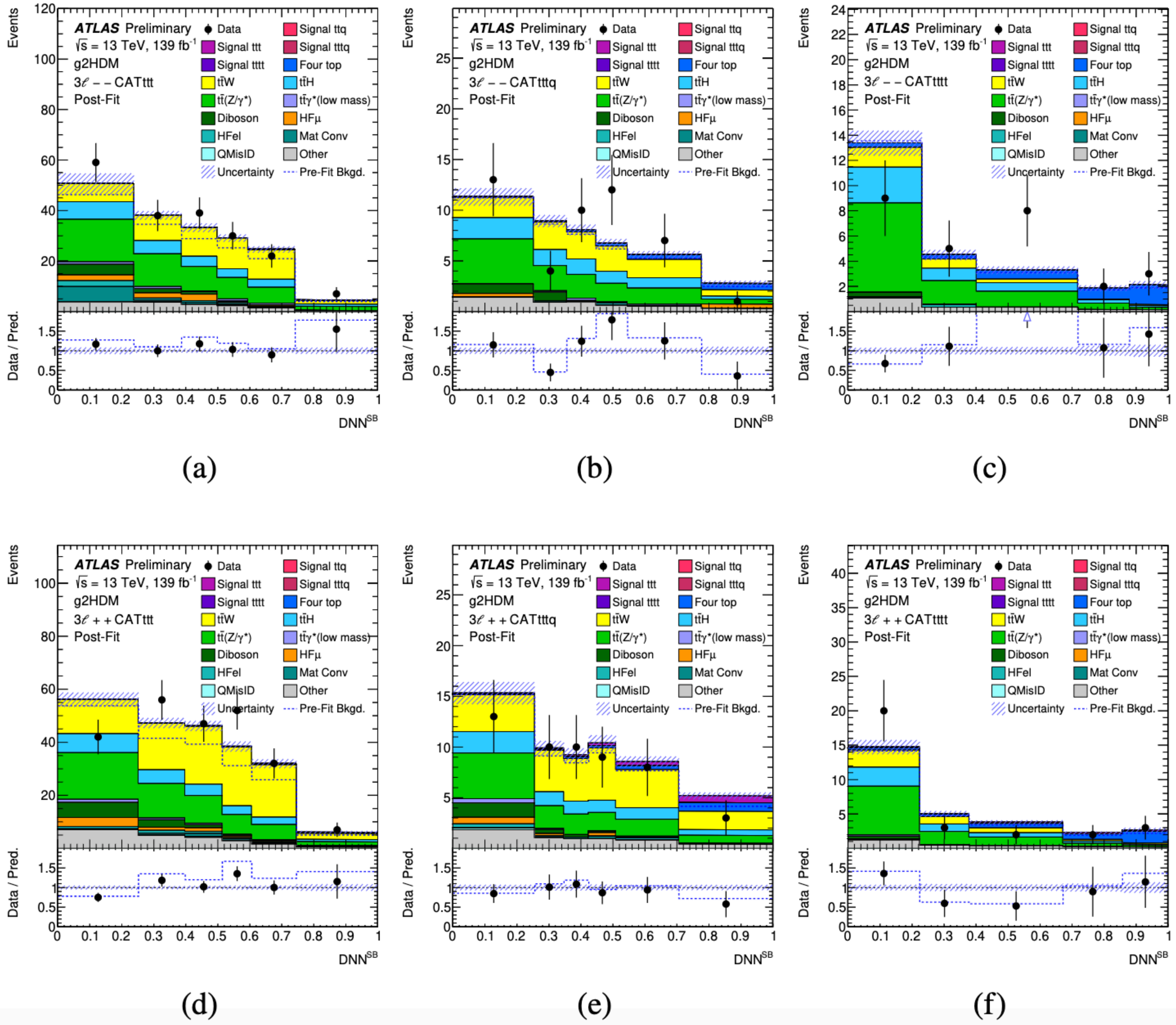
Backup

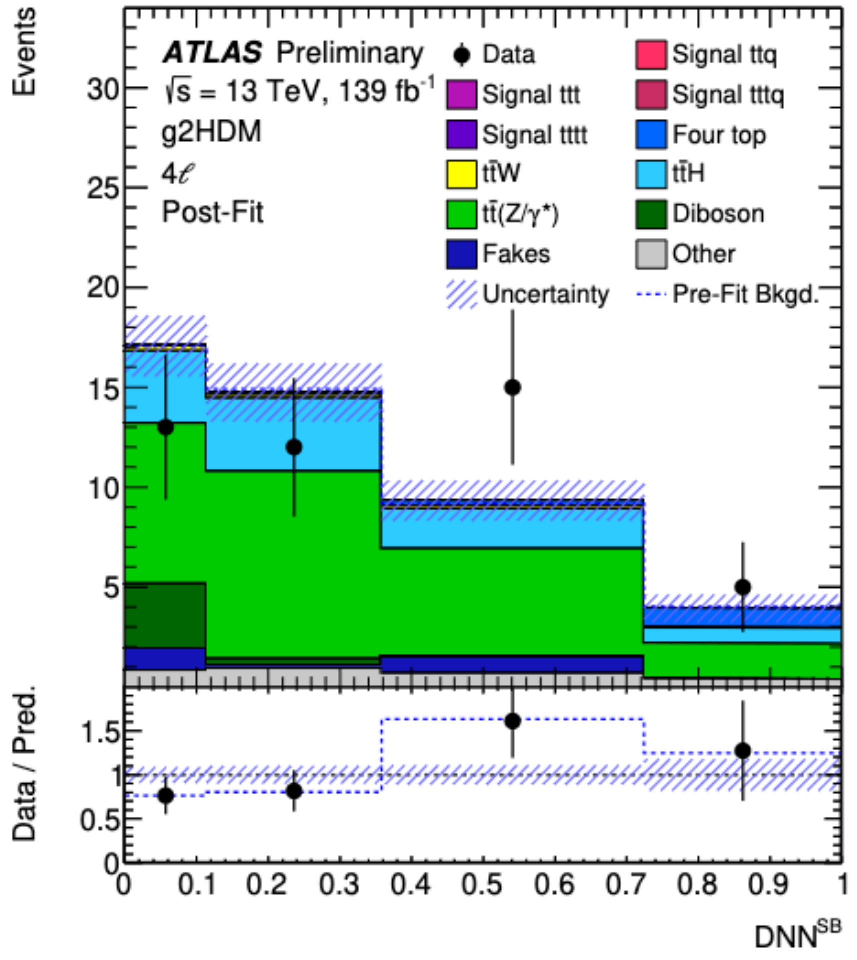


(a) (b) (c) (d)

(e) (f) (g) (h)

Backup


ATLAS Preliminary
√s = 13 TeV, 139 fb⁻¹
g2HDM
3ℓ − − CATttt
3ℓ − − CATtttq
3ℓ − − CATtttt
3ℓ + + CATttt
3ℓ + + CATtttq
3ℓ + + CATtttt
Post-Fit
Data
Signal ttq
Signal ttt
Signal tttq
Signal tttt
Four top
ttW
ttH
tt(Z/γ*)
ttγ*(low mass)
Diboson
HFμ
HFel
Mat Conv
QMisID
Other
Uncertainty
Pre-Fit Bkgd.
Events
Data / Pred.
DNN^SB
(a)
(b)
(c)
(d)
(e)
(f)


ATLAS Preliminary
√s = 13 TeV, 139 fb⁻¹
g2HDM
4ℓ
Post-Fit
Data
Signal ttq
Signal ttt
Signal tttq
Signal tttt
Four top
ttW
ttH
tt(Z/γ*)
Diboson
Fakes
Other
Uncertainty
Pre-Fit Bkgd.
Events
Data / Pred.
DNN^SB

Backup

$\hat{\lambda}_e^{\text{Mat Conv}} = 1.16 \pm 0.29$

$\hat{\lambda}_\mu^{\text{had}} = 0.92 \pm 0.18$

$\hat{\lambda}_e^{\text{had}} = 1.05 \pm 0.31$

$\hat{\lambda}_{t\bar{t}W} = 1.50 \pm 0.14$

$\hat{\lambda}_{VV} = 0.85 \pm 0.30$

$\hat{\lambda}_{t\bar{t}Z} = 0.97 \pm 0.19$

$\hat{\lambda}_{t\bar{t}Z} = 0.97 \pm 0.19$

The sample used to model the $t\bar{t}W$ ($t\bar{t}Z/\gamma^*$) background was generated using SHERPA-2.2.10 [79] (SHERPA-2.2.11), where the matrix element (ME) was calculated for up to one (zero) additional parton at next-to-leading-order (NLO) in QCD and up to two partons at LO in QCD using COMIX [80] and OPENLOOPS [81] and merged with the SHERPA parton shower (PS) [82] using the MEPS@NLO prescription [83], with a CKKW merging scale of 30 GeV for the $t\bar{t}W$ sample. These samples are generated using the `NNPDF3.0nnlo` [84] PDF set. Both the factorisation and renormalisation scales are set to $\mu_r = \mu_f = m_T/2$ in the $t\bar{t}W$ sample, where m_T is defined as the scalar sum of the transverse masses $\sqrt{m^2 + p_T^2}$ of the particles generated from the ME calculation. The LO $t\bar{t}W$ electroweak (EW) contributions are obtained from a dedicated sample simulated with SHERPA-2.2.10. The invariant mass of the lepton pair ($m_{\ell^+\ell^-}$) in the $t\bar{t}Z/\gamma^*$ sample is set to be greater than 1 GeV. The complete $t\bar{t}W$ simulation is normalised to the total cross section of $\sigma(t\bar{t}W) = 614.7$ fb that comes from the SHERPA configuration outlined above considering NLO QCD and NLO EWK effects, based on a similar strategy as used in Ref. [85]. The $t\bar{t}Z/\gamma^*$ sample is normalised to the cross section $\sigma(t\bar{t}Z/\gamma^*) = 839$ fb, calculated at NLO QCD and NLO EW accuracy using MADGRAPH5_AMC@NLO [86] and scaled by an off-shell correction estimated at one-loop level in α_s.