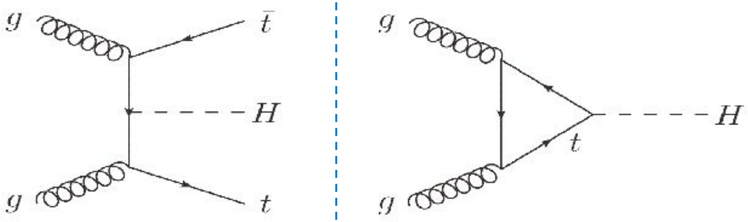
$t\bar{t}H$ in the multilepton channel

Vera Maiboroda
CEA Saclay / Université Paris-Saclay

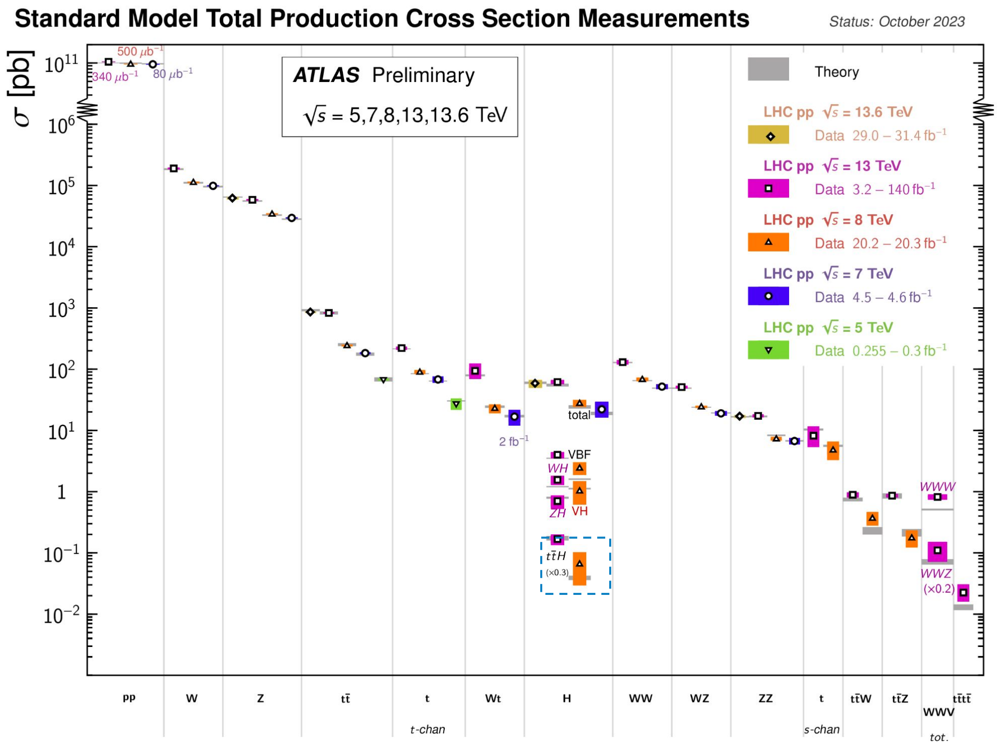
On behalf of Saclay ATLAS team

Top Yukawa-coupling via $t\bar{t}H$

- Top Yukawa coupling (y_t) is expected to be **of order one**
  → idea about the scale of New Physics
- $t\bar{t}H/tH$ production cross-section measurement is the only **direct** way to measure y_t
- $t\bar{t}H$ allows to probe the **CP** structure of the top-Higgs coupling
  (talk by Alberto Carnelli)



direct *indirect*



$t\bar{t}H$ channels

- Production of $t\bar{t}H$ accounts for about 1% of the total rate
- Covering as many decay channels as possible:

$H\rightarrow\gamma\gamma$	ATLAS, CMS
$H\rightarrow b\bar{b}$	ATLAS, CMS
$H\rightarrow ZZ\rightarrow 4\ell$	ATLAS, CMS
$H\rightarrow\tau\tau$	ATLAS

- $t\bar{t}H$ in multilepton final state: clean final state with leptons, moderate irreducible background

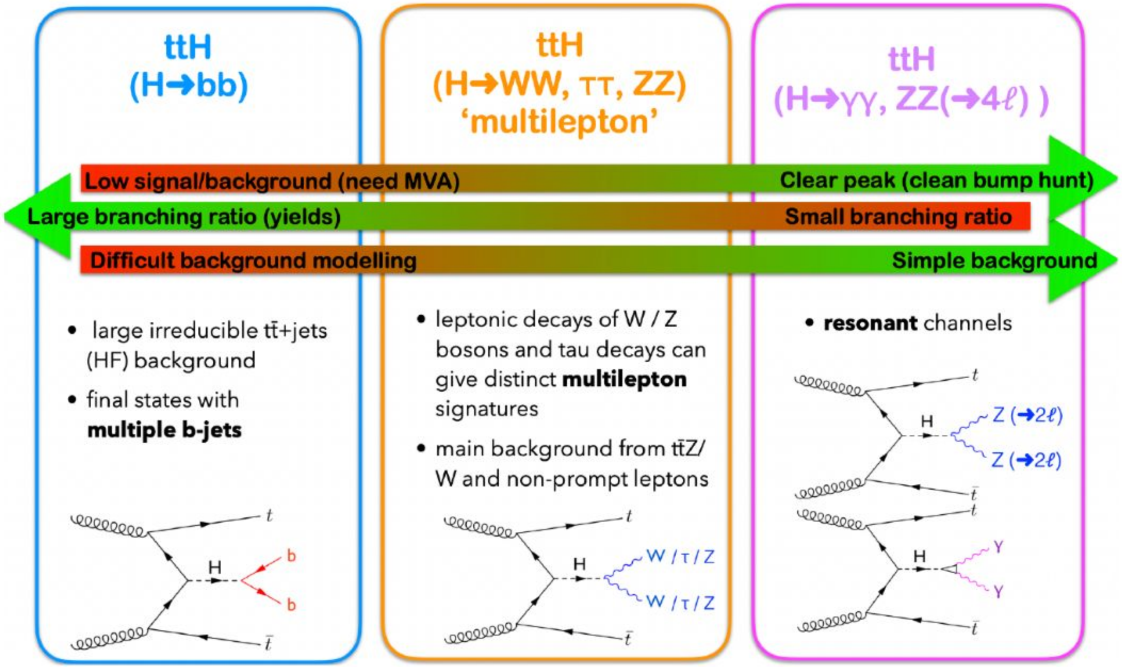
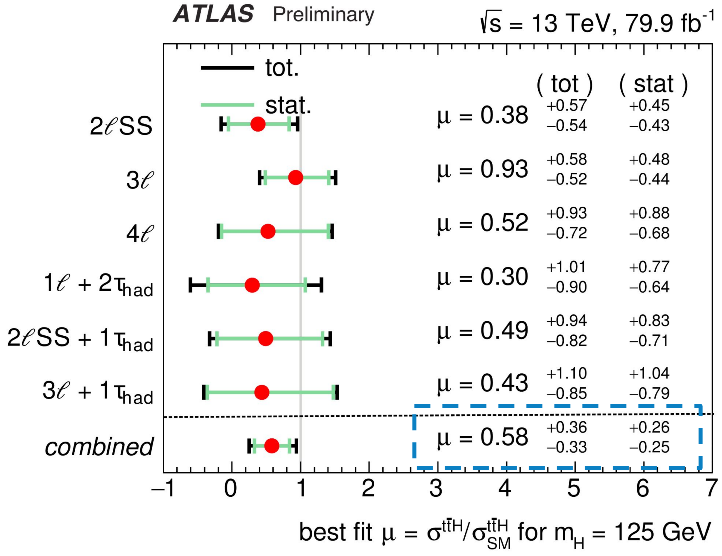
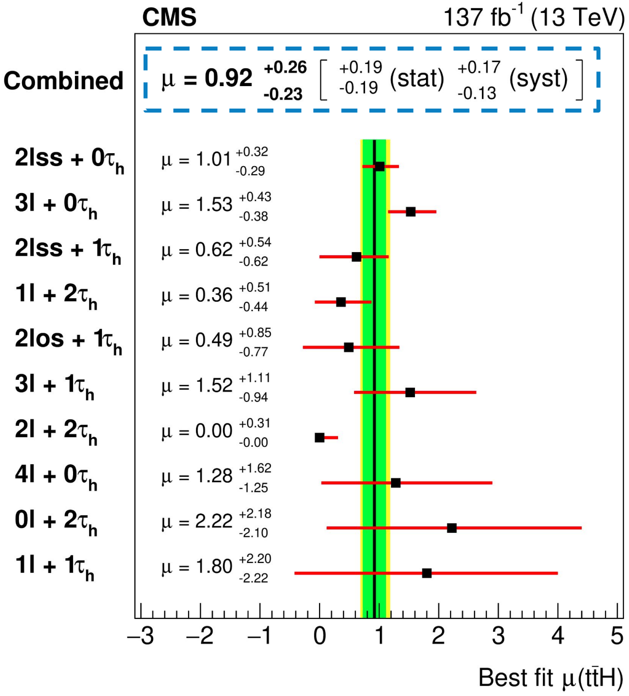


Figure from Tamara Vazquez Schroeder

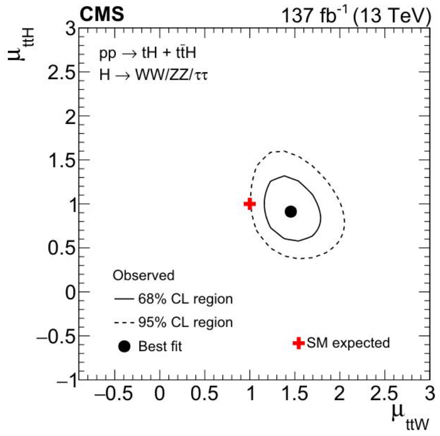
Current results $t\bar{t}H$

- $t\bar{t}H$ observation by ATLAS and CMS in 2018 with partial Run 2 dataset
- $t\bar{t}H$ in multilepton by ATLAS (80 fb^{-1}) and CMS (137 fb^{-1})
  ATLAS measured $\mu(t\bar{t}H) \approx 0.58$ with an observed significance of 1.8σ



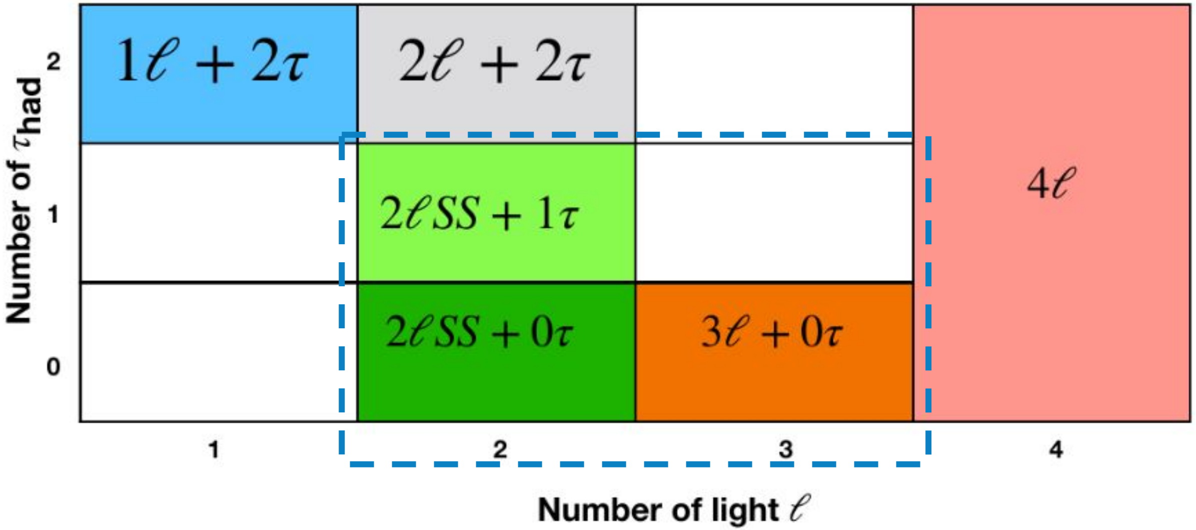


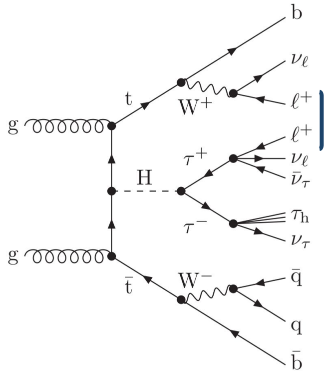
Observed $t\bar{t}W$ mismodelling



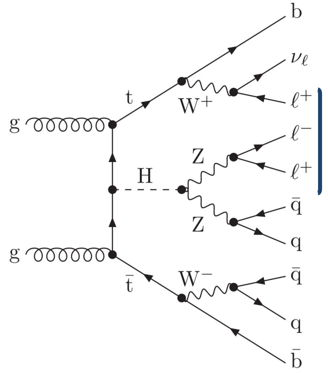
Analysis setup and strategy

- Full run 2 dataset (140 fb^{-1})
- $t\bar{t}H$ ML inclusive cross-section measurement including 6 different channels
- Simplified template cross-section (STXS) measurement
- Higgs CP interpretation





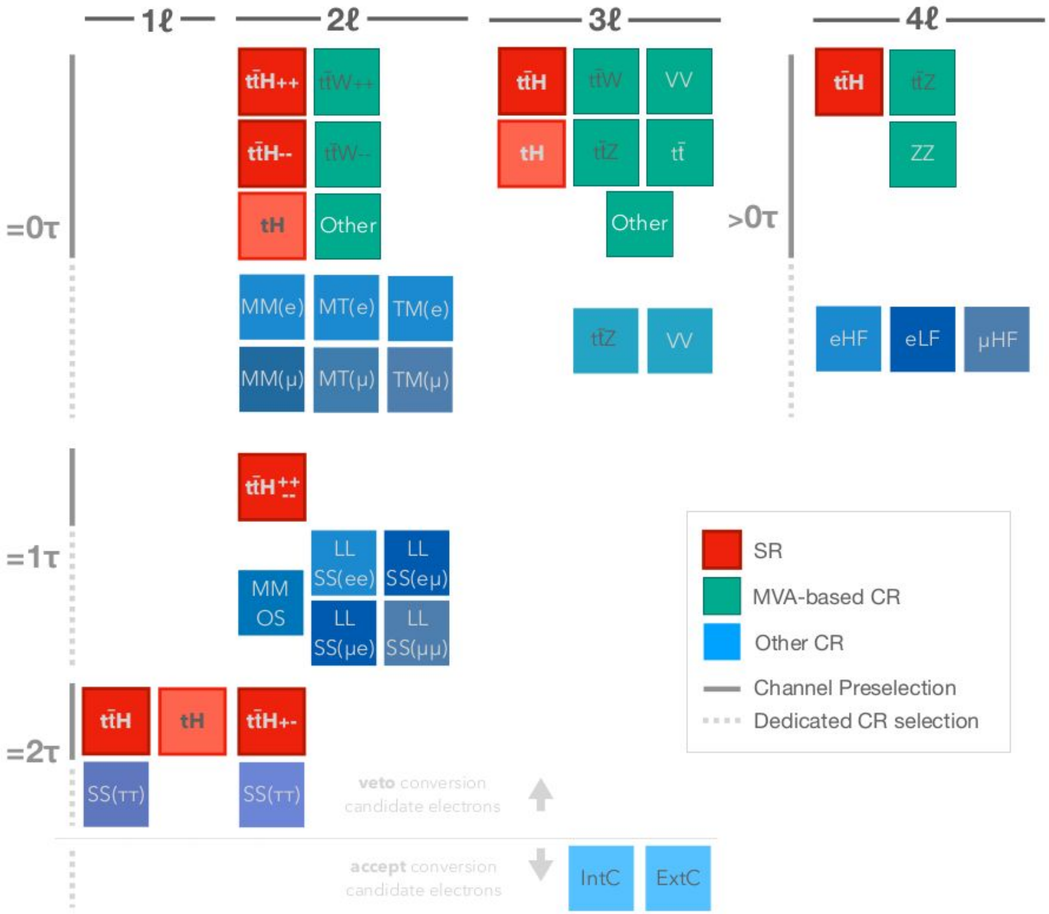
2ℓSS final state



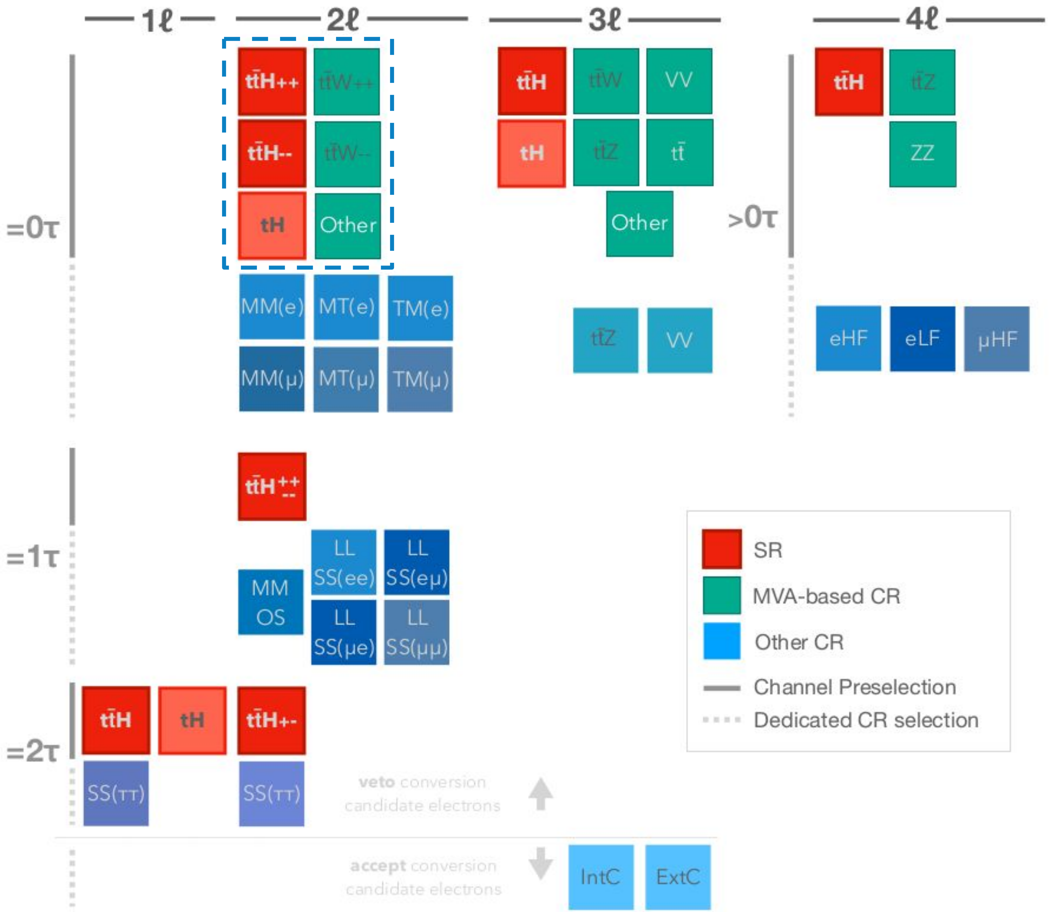
3ℓ final state

Regions definition

- Signal regions (SR)
- MVA-based control regions (CR)
- "Other" CRs
  from the template fit approaches used to model fake lepton backgrounds



Regions: 2ℓSS + 0τ



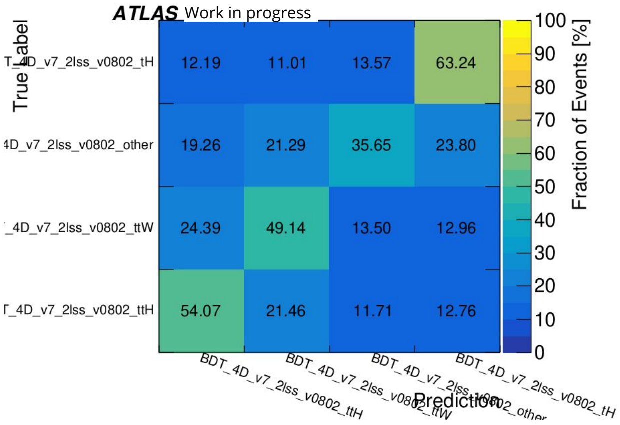
Regions: 2ℓSS + 0τ

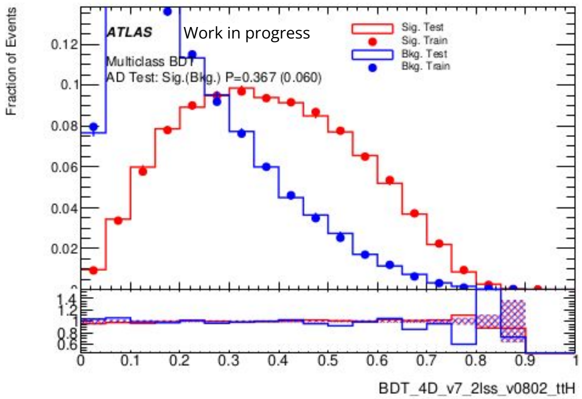
- Signal and MVA-based control regions definition
  - Pre-MVA selection:

	$2\ell SS{+}0\tau_{had}$
τ candidates	==0 M
Leptons counting	==2 T: $p_T > 15$ GeV
Lepton details	SS
njets	≥ 3
nbjets (at 85% WP)	≥ 1

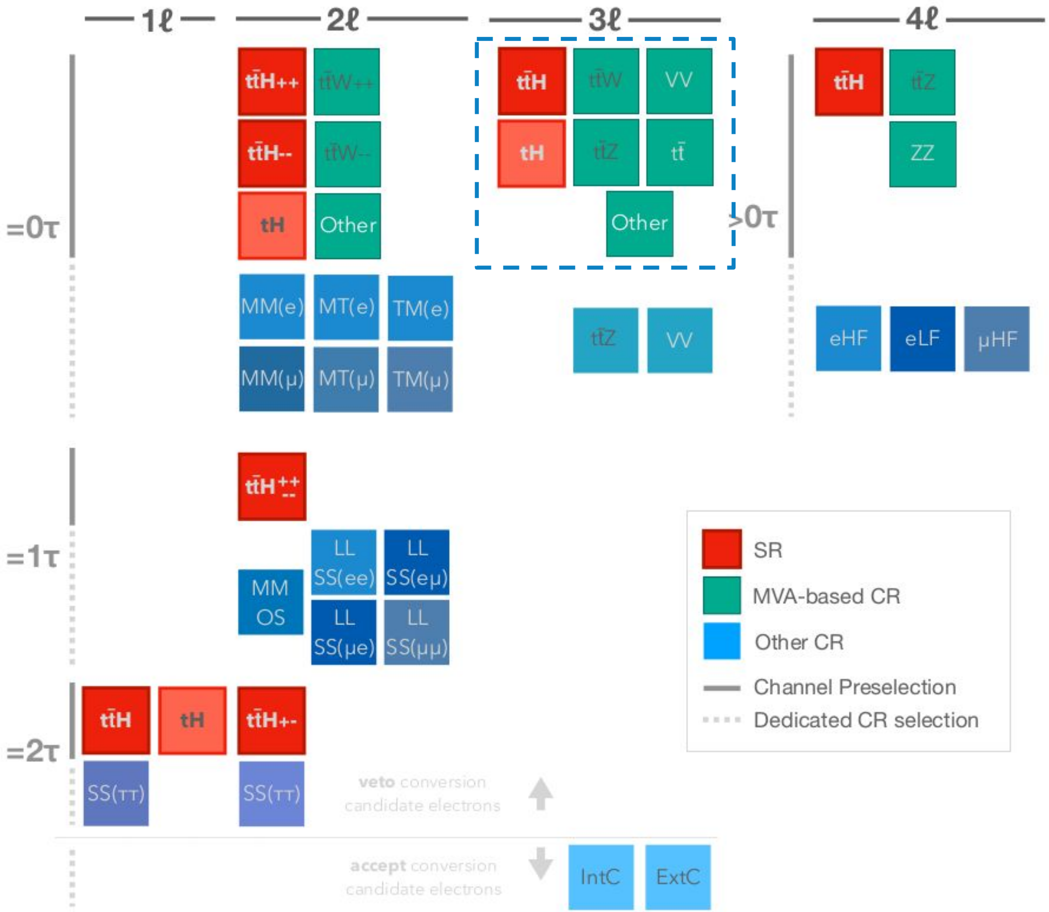
  - MVA selection:
    - Multi-class BDT: *tt̄H, tH, tt̄W* and *Other* as output nodes
    - 20 variables to train: NJets, m_{ll}, dR (l, jet), etc.
    - Regions are defined based on the highest BDT score

- Dominant prompt backgrounds: *tt̄W, tt̄Z*





Regions: 3ℓ + 0τ



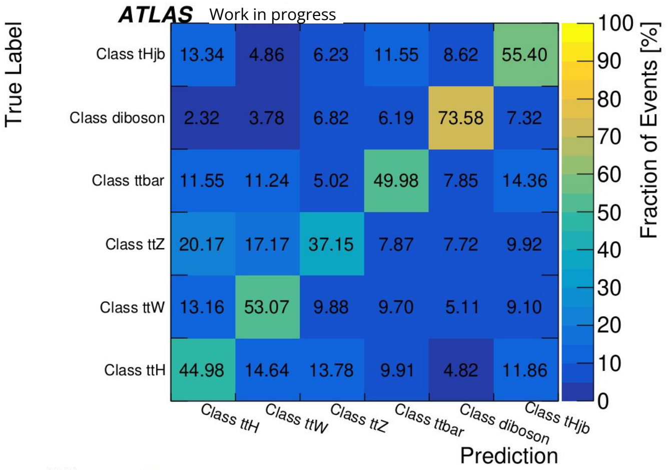
Regions: 3ℓ + 0τ

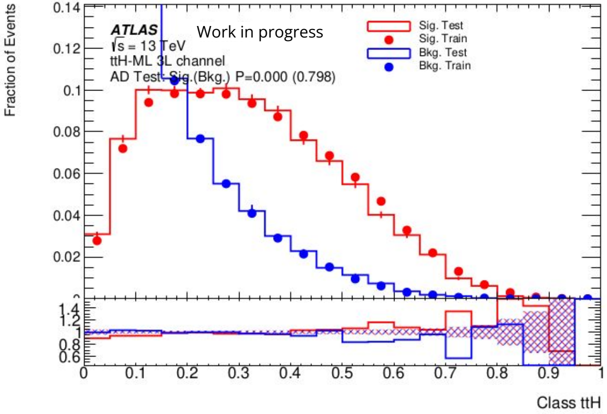
- Signal and MVA-based control regions definition
  - Pre-MVA selection:

	$3\ell+0\tau_{had}$
τ candidates	==0 M
Leptons counting	==3 (T,T,L): $p_T > 15, 15, 10$ GeV
Lepton details	OS (to others): L $p_T > 10$ GeV SS pair: T $p_T > 15$ GeV OS pair: $\lvert m(ll) - m_Z \rvert > 10$ GeV and $m(ll) > 12$ GeV
njets	≥ 2
nbjets (at 85% WP)	≥ 1

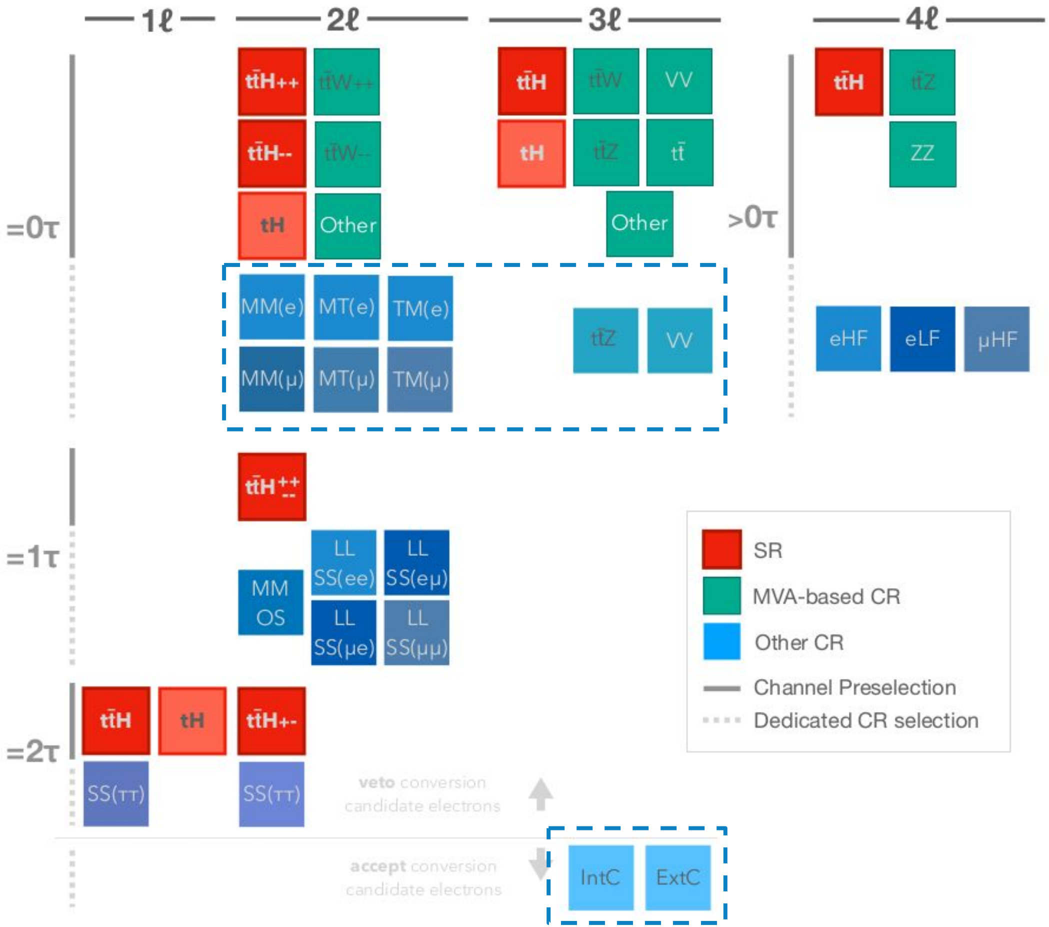
  - MVA selection:
    - Multi-class BDT: *tt̄H, tH, tt̄W, tt̄Z, VV* and *tt̄* as output nodes
    - 25 variables to train: NJets, m_{ll}, dR (l, jet), etc.
    - Regions are defined by optimized cuts on BDT scores

- Dominant prompt backgrounds: *tt̄W, tt̄Z*



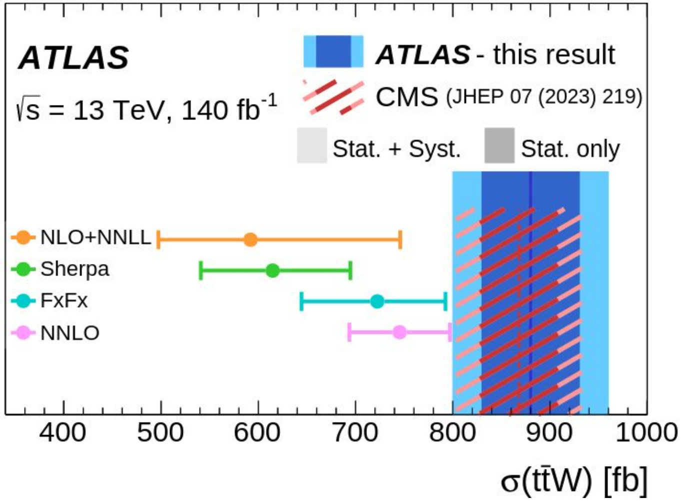


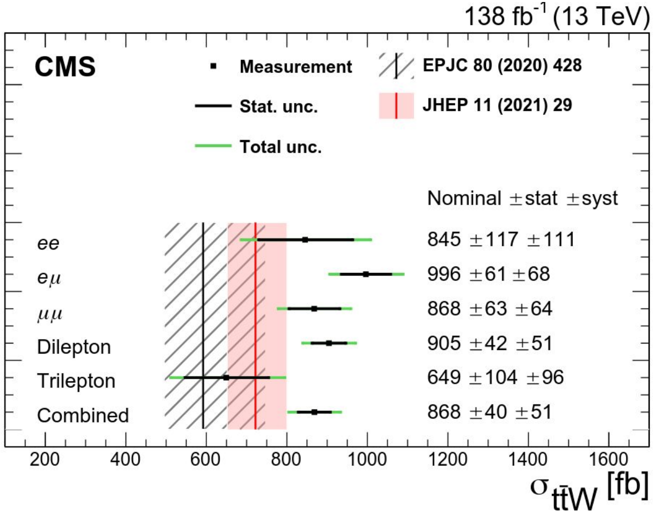
Control regions definitions: 2ℓSS + 0τ and 3ℓ + 0τ

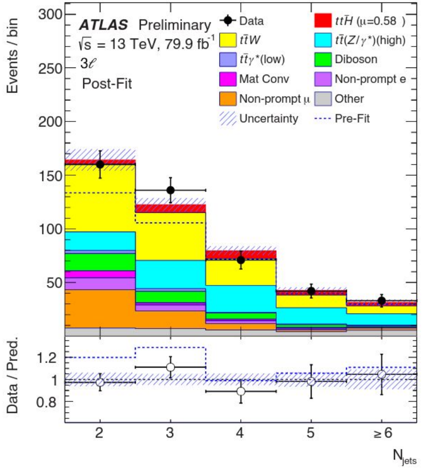


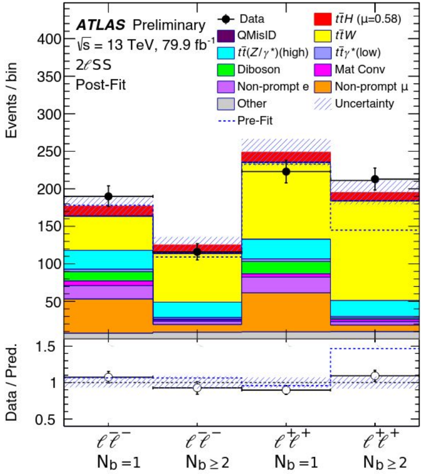
Mismodelled main background: $t\bar{t}W$

- $t\bar{t}W$ is the main background for $t\bar{t}H$ multilepton
- Observed $t\bar{t}W$ mismodelling is **20-50% larger** cross section than predicted
  - Theory: NLO + NNL: 606 fb ± 7%
    NLO + FxFx: 722 fb ± 10%
  - Experiment: ATLAS: 880 fb ± 6% (stat) ± 8% (syst)
    CMS: 868 fb ± 5% (stat) ± 6% (syst)

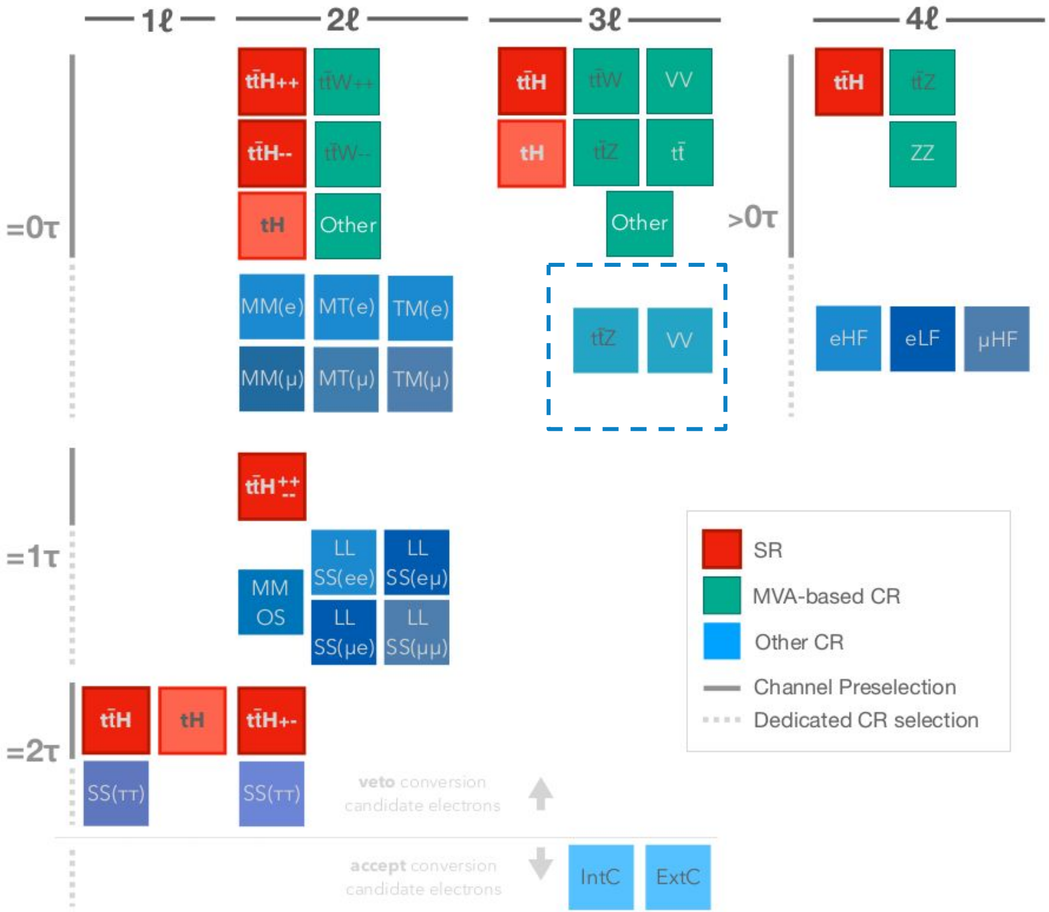








Control regions definitions: $3\ell + 0\tau$



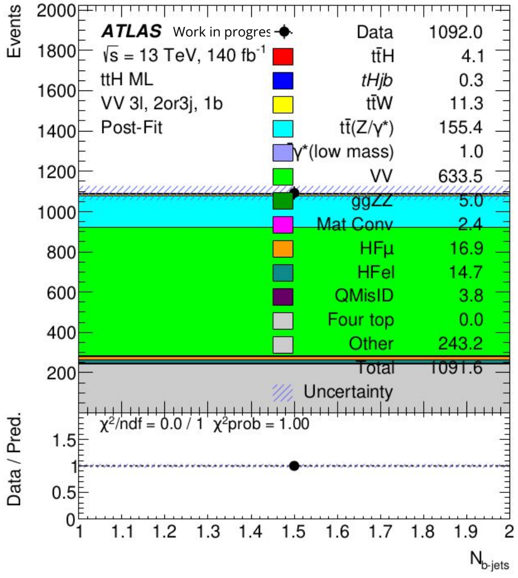
Backgrounds

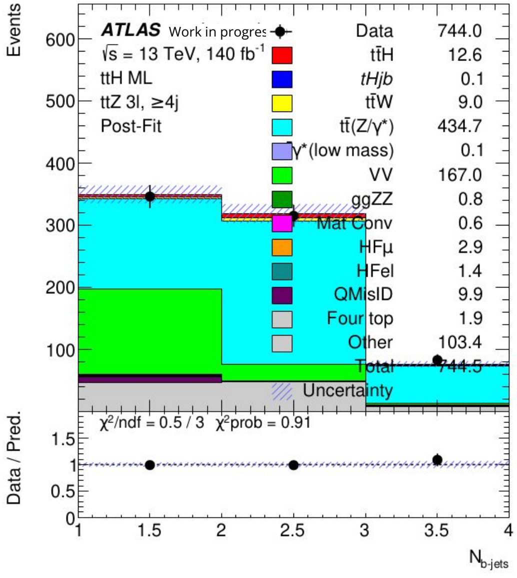
Irreducible background:

- Events with **prompt leptons** produced in
  - W/Z boson decays
  - leptonic τ-lepton decays
  - internal conversions
- Source:
  - Main: $t\bar{t}W$, $t\bar{t}Z/\gamma^*$ and VV
  - Smaller: tZ, tW, tWZ, $t\bar{t}WW$, VVV, $t\bar{t}t\bar{t}$, $t\bar{t}t$
- Estimated from **simulation**

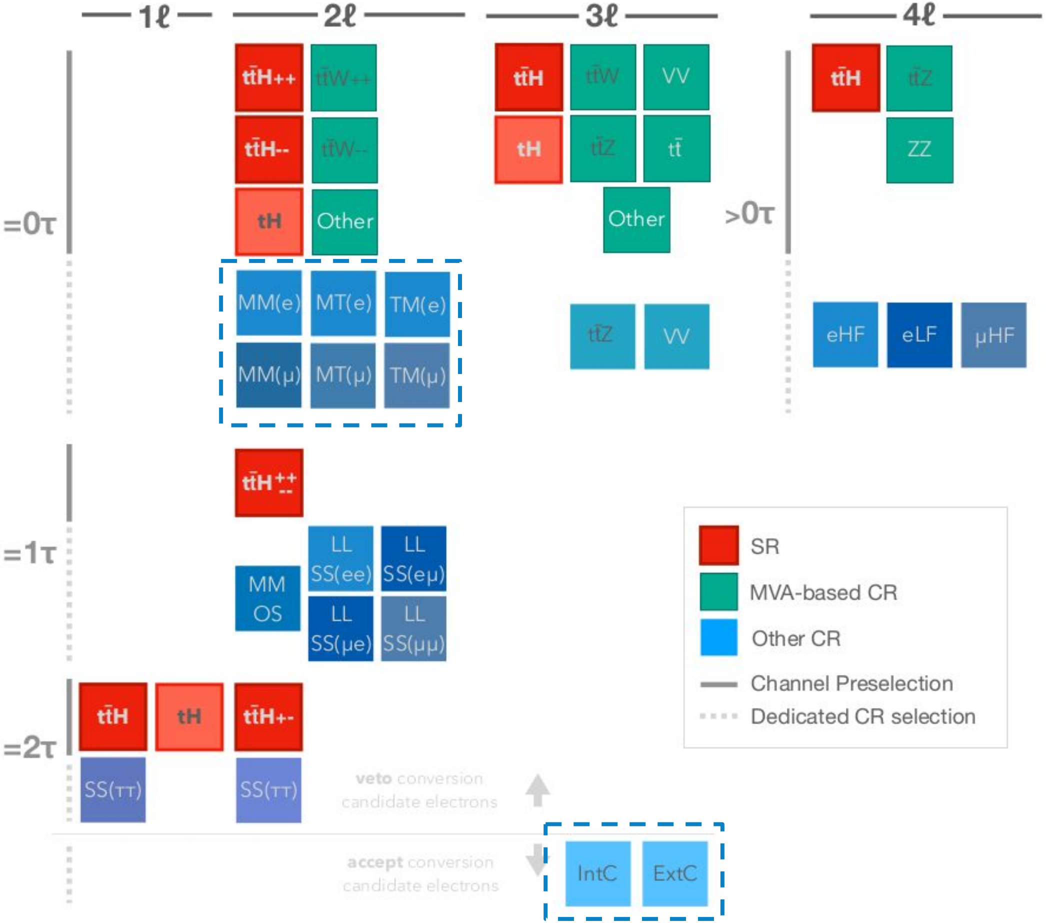
Examples:

- Jet multiplicity not well modelled in VV sample: N_{jets}-dependent data-driven corrections (derived in $t\bar{t}W$)
- Cut-based $3l$ VV and $t\bar{t}Z$ CRs taken from $t\bar{t}W$ analysis





Control regions definitions: 2ℓSS + 0τ

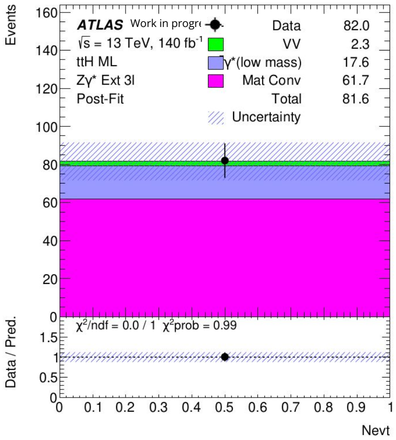


Backgrounds

Reducible background:

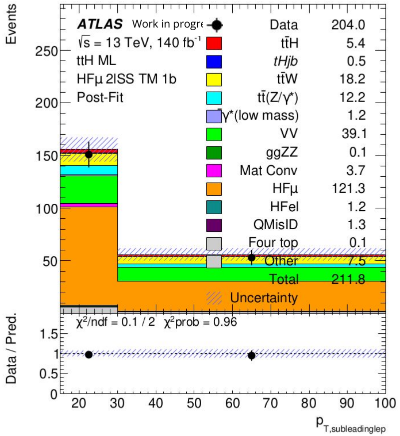
- Events with at least
  - one prompt charge-flip electron
  - one fake lepton
- Source:
  - material conversions
  - hadron decays
  - improper reconstruction of other particles
- Estimated from **simulation** mainly using template method

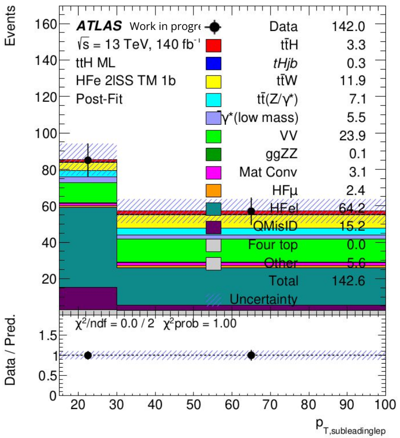
Backgrounds



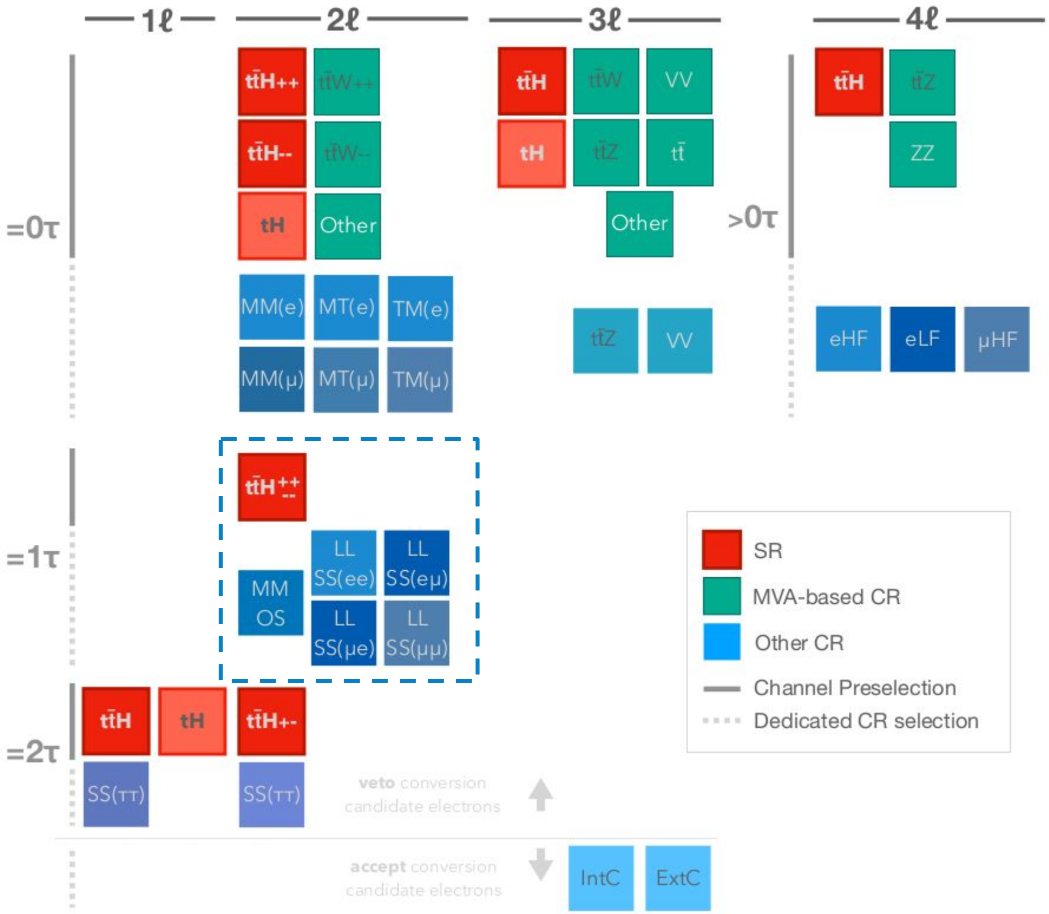
Examples:

- Fake/non-prompt lepton: estimated from simulation, with the normalisation determined by the likelihood fit.
  - 2 conversion CRs (internal and external): *Z(→ μμ)γ*
  - 6 2lSS CRs enriched in contributions from HF fakes in $t\bar{t}$ events
    - split by flavour of subleading lepton
    - split by prompt leptons veto working points
  - 4 NFs: HF *e* and HF *μ* fakes, material and internal conversion



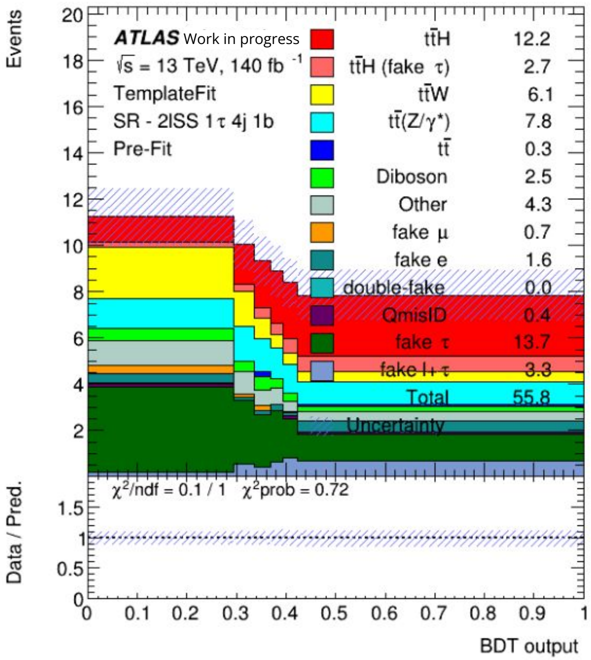


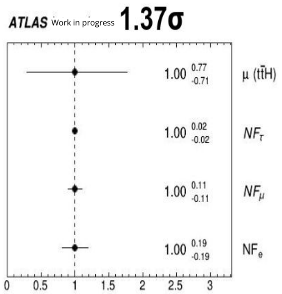
Regions definition: 2ℓSS + 1τ



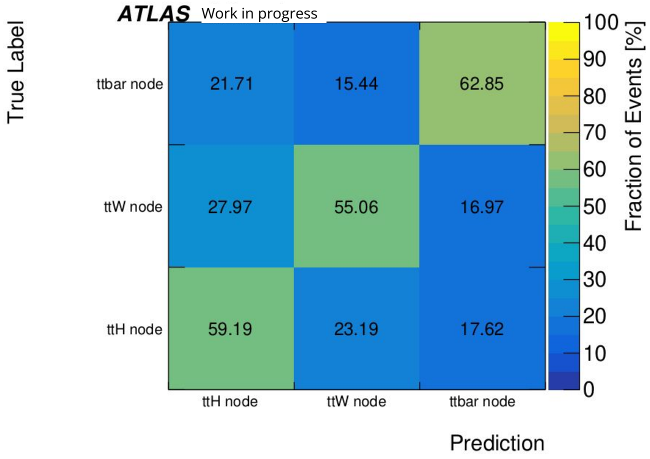
Regions definition: 2ℓSS + 1τ

- Similar BDT-based procedure, but used to define the SR only
- Dominant background from fakes, no *ttW* or *ttZ* CRs



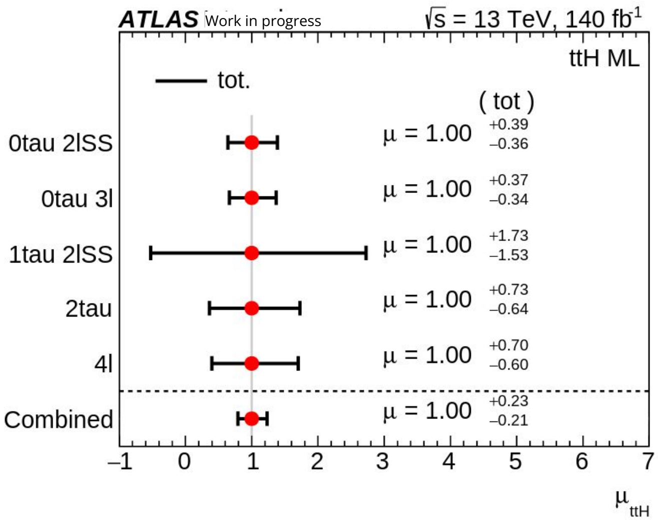


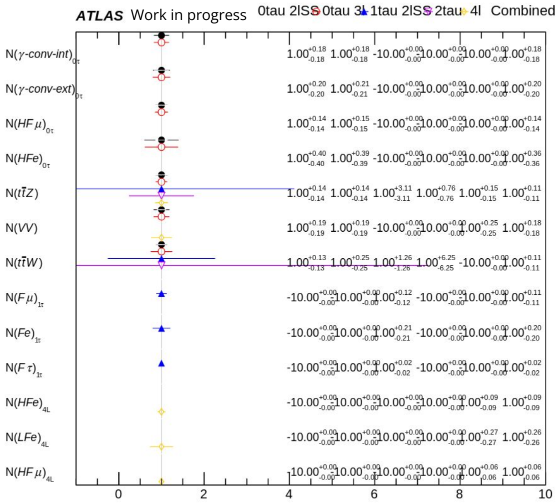
No ttW or ttZ NF in 2lSS+1tau channel

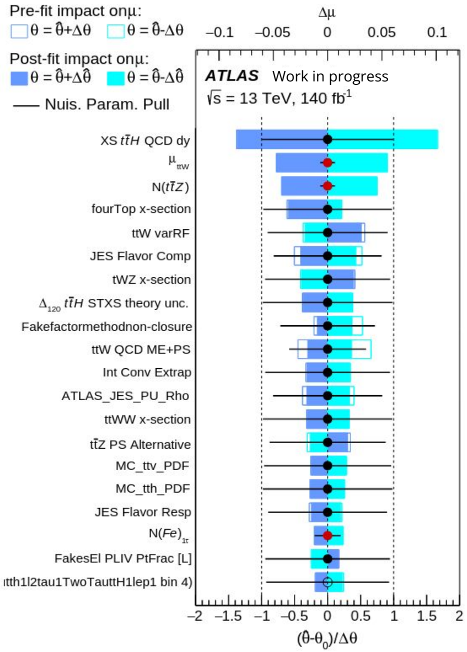


Results combined: Asimov

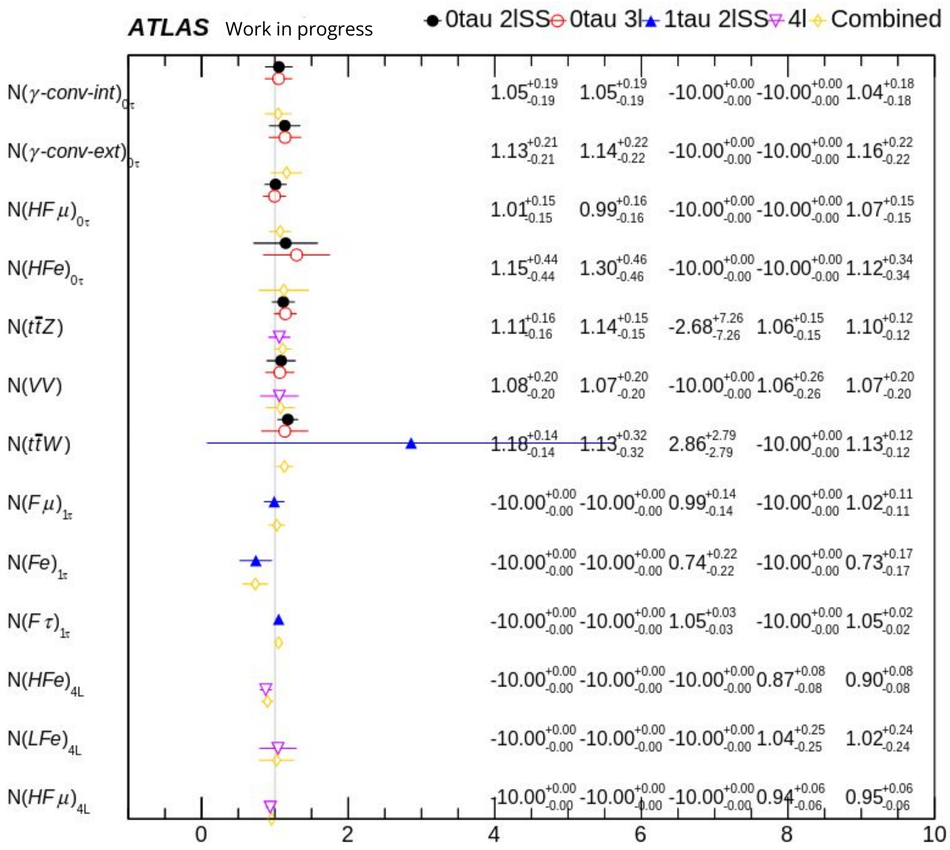
- Expected significance is 5.7σ (3.1σ in previous analysis)
- Measure *ttH* ML with significance of 5.4σ (not all systematics are included yet)
- Main systematics: *ttH* cross-section uncertainty, *ttW* and *ttZ* NFs





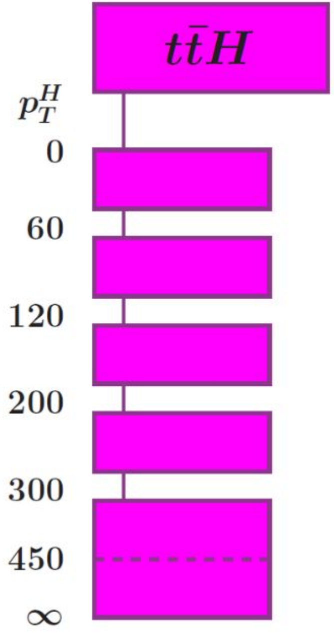


Results combined: CR only with data



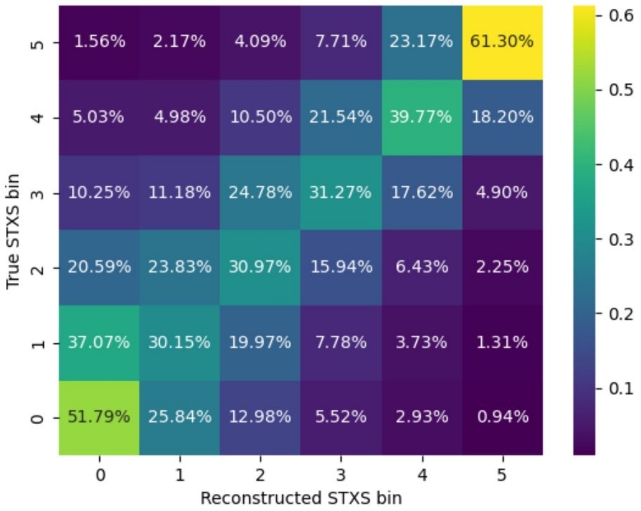
STXS setup

- Simplified template cross-section (STXS)
  - SR splitting of Higgs production processes on Higgs pT
  - Simplify combination between channels/measurements
  - Isolate possible BSM effects,
    minimize the dependency on theory uncertainties
- 6 bins in H_{pT} in STXS framework for $t\bar{t}H$

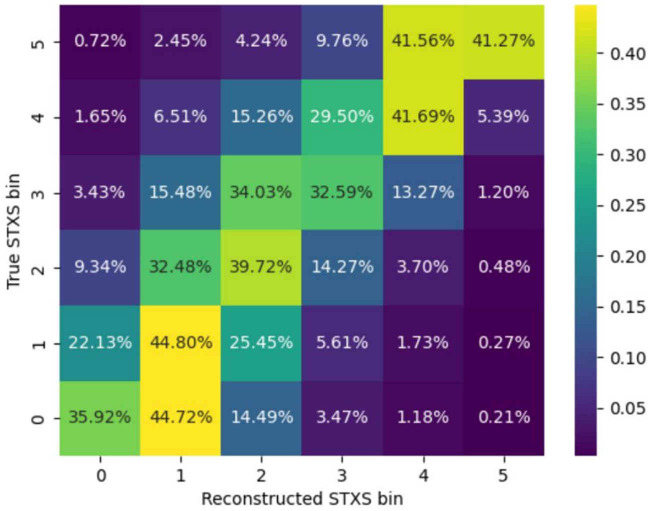


STXS setup

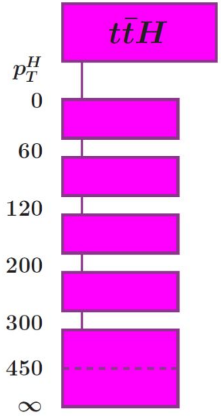
- 6 bins in H_{pT} in STXS framework for $t\bar{t}H$
- H_{pT} difficult to reconstruct:
  many neutrinos in *H* decay and not reconstructed soft particles
- Use GNN for H_{pT} reconstruction

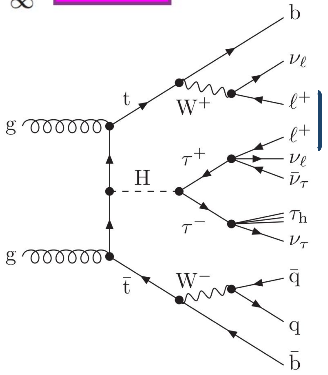


2ℓSS+0τ channel



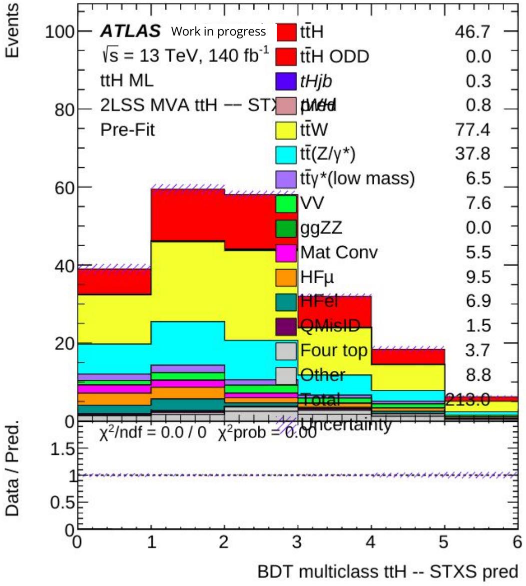
3ℓ+0τ channel

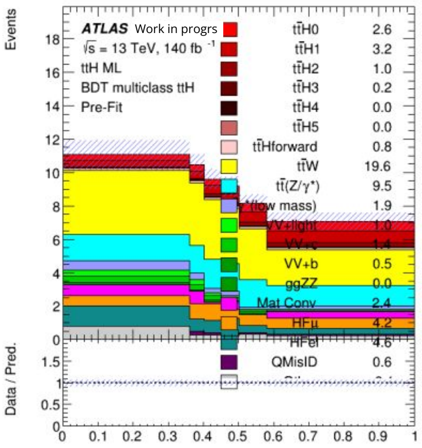


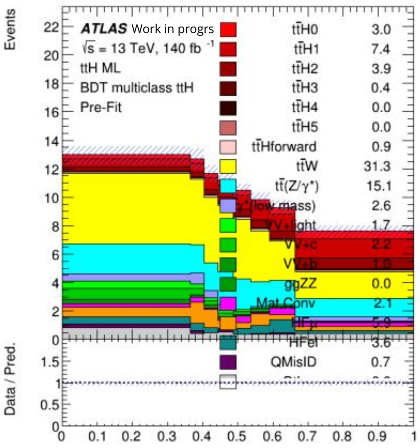


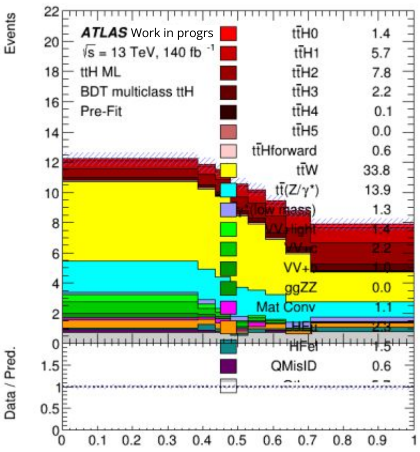
2ℓSS final state

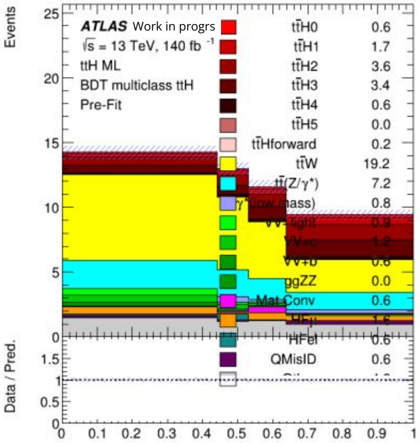
STXS 2ℓSS + 0τ - - for the STXS fit setup

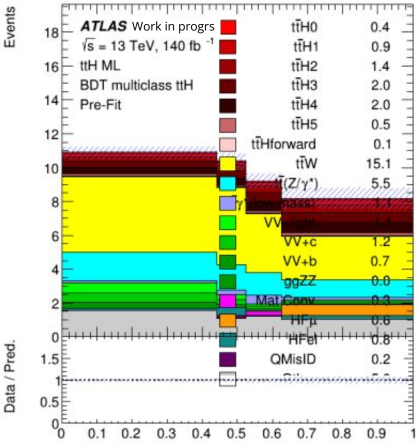








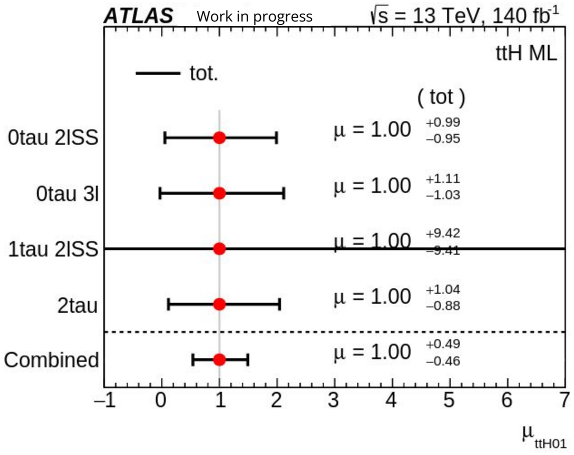




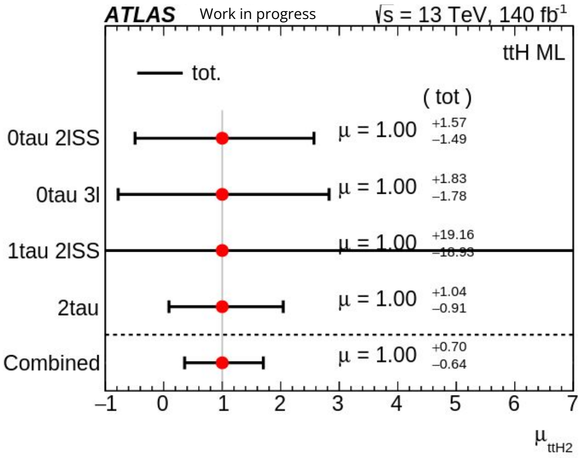
STXS results combined Asimov

- Sensitivity to STXS bins not strong enough to measure all 6 bins
- Significant improvement in combined fit compared to single channels

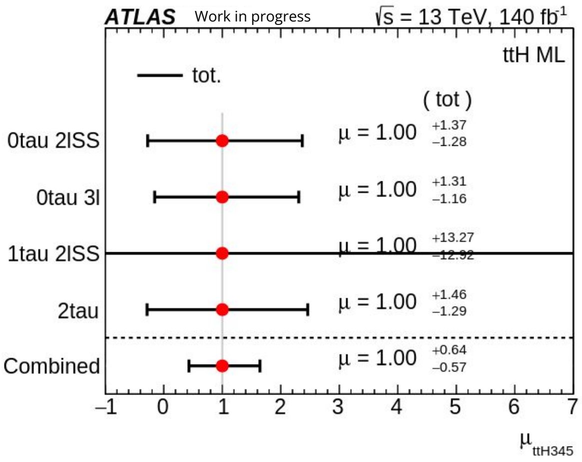
$p_T^H < 120$ GeV



120 GeV $\leq p_T^H < 200$ GeV



200 GeV $\leq p_T^H$



Conclusions

- Still some work on harmonization and including all systematics is to be done
- Wait for the unblinded results

Thanks for the attention

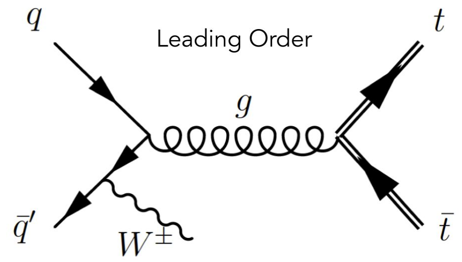
Monte Carlo samples

The configurations used for event generation of signal and background processes.

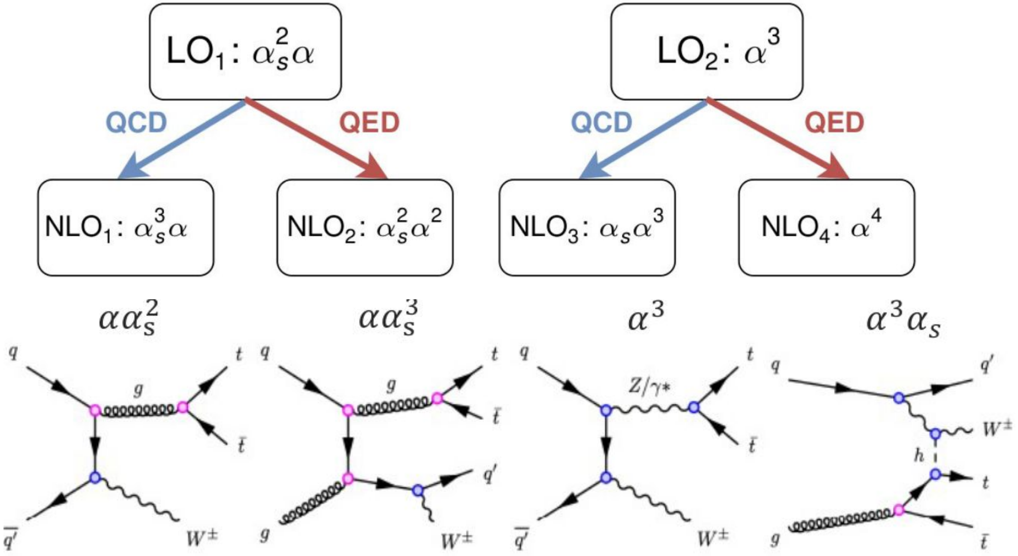
Systematics samples are in grey.

Process	Generator	ME order	Parton shower	PDF	Tune
$t\bar{t}H$	Powheg-BOX	NLO	Pythia 8	NNPDF3.0nlo	A14
	(Powheg-BOX)	(NLO)	(Herwig7.0.4)	(NNPDF3.0nlo)	(H7-UE-MMHT)
	(MG5_aMC)	(NLO)	(Pythia 8)	(NNPDF3.0nlo)	(A14)
$t\bar{t}W$	Sherpa 2.2.10	MePs@Nlo	Sherpa	NNPDF3.0nnlo	Sherpa default
	(MG5_aMC)	(FxFx NLO)	(Pythia 8)	(NNPDF3.0nlo)	(A14)
	(Powheg)	(NLO)	(Pythia 8)	(NNPDF3.0nlo)	(A14)
	(Powheg)	(NLO)	(Herwig 7)	(NNPDF3.0nlo)	(H7-UE-MMHT)
$t\bar{t}W$ (EW)	Sherpa 2.2.10	LO	Sherpa	NNPDF3.0nnlo	Sherpa default
	(MG5_aMC)	(LO)	(Pythia 8)	(NNPDF3.0nlo)	(A14)
$t\bar{t}\ell\ell$	MG5_aMC	NLO	Pythia 8	NNPDF3.0nlo	A14
	(MG5_aMC)	(NLO)	(Herwig 7)	(NNPDF3.0nlo)	(H7-UE-MMHT)
	(MG5_aMC)	(NLO)	(Pythia 8)	(NNPDF3.0nlo)	(A14 Var3c)
$t\bar{t} \to W^+bW^-\bar{b}\ell^+\ell^-$	MG5_aMC	LO	Pythia 8	NNPDF3.0lo	A14
$t\bar{t}t\bar{t}$	MG5_aMC	NLO	Pythia 8	NNPDF3.1nlo	A14
$t\bar{t}$	Powheg-BOX	NLO	Pythia 8	NNPDF3.0nlo	A14
	(Powheg-BOX)	NLO	(Herwig7.1.3)	(NNPDF3.0nlo)	(H7-UE-MMHT)
$t\bar{t}t$	MG5_aMC	LO	Pythia 8	NNPDF2.3lo	A14
Single top (t-, Wt-, s-channel)	Powheg-Box	NLO	Pythia 8	NNPDF3.0nlo	A14
VV, $qqVV$, VVV	Sherpa 2.2.2(1)	MePs@Nlo	Sherpa	NNPDF3.0nnlo	Sherpa default
$Z \to \ell^+\ell^-$	Sherpa 2.2.1	MePs@Nlo	Sherpa	NNPDF3.0nnlo	Sherpa default
$Z \to \ell^+\ell^-(\gamma \to e^+e^-)$	Powheg-BOX	NLO	Pythia 8	CTEQ6L1nlo	A14
$Z \to \ell^+\ell^-(\gamma* \to e^+e^-)$	Powheg-BOX	NLO	Pythia 8	CTEQ6L1nlo	A14

Monte Carlo samples: ttW



Process	Generator	ME order	Parton shower	PDF	Tune
$t\bar{t}H$	POWHEG-BOX	NLO	PYTHIA 8	NNPDF3.0NLO	A14
	(POWHEG-BOX)	(NLO)	(HERWIG7.0.4)	(NNPDF3.0NLO)	(H7-UE-MMHT)
	(MG5_aMC)	(NLO)	(PYTHIA 8)	(NNPDF3.0NLO)	(A14)
$t\bar{t}W$	SHERPA 2.2.10	MEPS@NLO	SHERPA	NNPDF3.0NNLO	SHERPA default
	(MG5_aMC)	(FxFx NLO)	(PYTHIA 8)	(NNPDF3.0NLO)	(A14)
	(POWHEG)	(NLO)	(PYTHIA 8)	(NNPDF3.0NLO)	(A14)
	(POWHEG)	(NLO)	(HERWIG 7)	(NNPDF3.0NLO)	(H7-UE-MMHT)



Nominal MC sample: Sherpa

Alternative samples: MG5_aMC, Powheg

Selection criteria applied for the pre-selection regions

	2ℓSS+$0\tau_{\text{had}}$	3ℓ+$0\tau_{\text{had}}$	4ℓ
τ_{had} candidates	==0 M	==0 M	–
Leptons counting	==2 T: p_{T} >15 GeV	==3 (T,T,L): p_{T} > 15, 15, 10 GeV	==4 L: p_{T} > 10 GeV
Lepton details	SS	OS (to others): L p_{T} > 10 GeV SS pair: T p_{T} >15 GeV OS pair: $\|m(ll) - m_Z\| > 10$ GeV and $m(ll) > 12$ GeV	Sum charge = 0 OS pairs: $m(ll) > 12$ GeV $\|m(llll) - m_H\| > 5$ GeV
N_{jets}	≥ 3	≥ 2	≥ 2
$N_{b-\text{jets}}$ (@ 85% WP)	≥ 1	≥ 1	≥ 1

	2ℓSS+$1\tau_{\text{had}}$	1ℓ+$2\tau_{\text{had}}$	2ℓOS+$2\tau_{\text{had}}$
τ_{had} candidates	==1 M: p_{T} > 20 GeV	==2 OS M p_{T} > 20 GeV	==2 OS M p_{T} > 20 GeV
Leptons counting	==2 M: p_{T} >15 GeV	==1 L p_{T} > 27 GeV	==2 OS L p_{T} > 10 GeV
Lepton details	SS $\|m(ll) - m_Z\| > 10$		OS pair: $\|m(ll) - m_Z\| > 10$ GeV and $m(ll) > 12$ GeV
N_{jets}	≥ 3	≥ 3	–
$N_{b-\text{jets}}$	≥ 1 (@ 85% WP)	≥ 1 (@ 77% WP)	> 0 (@ 77% WP)

Loose, Medium, Medium exclusive, and Tight light lepton definitions

	e					μ			
	L	L'	M	M_{ex}	T	L	M	M_{ex}	T
LooseVar_Rad isolation	Yes					Yes			
Non-prompt lepton BDT (PLIV)	No		*Tight*	*Tight-not-VeryTight*	*VeryTight*	No	*Tight*	*Tight-not-VeryTight*	*VeryTight*
Identification	Loose	Tight				Loose	Medium		
Charge mis-assignment veto (ECIDS)	No	Yes				N/A			
Conversion rejection	No	Yes				N/A			
Transverse impact parameter significance $\lvert d_0 \rvert / \sigma_{d_0}$	< 5					< 3			
Lonzgitudinal impact parameter $\lvert z_0 \sin\theta \rvert$	< 0.5 mm								

Light lepton definitions

Channel	Cut-based Control Regions	Signal and MVA-based Control regions
2ℓSS	TM$_{ex}$,M$_{ex}$T, M$_{ex}$M$_{ex}$	TT
3ℓ	L'MM/LMM (L' for μ and L for e)	LTT (L for ℓ_0)
4ℓ	LLLL	
2ℓSS+1$\tau_{\rm had}$	L'L' and MM (for fake $\tau_{\rm had}$ CR)	MM
1ℓ+2$\tau_{\rm had}$ and 2ℓ+2$\tau_{\rm had}$	L	

Variables used for training the BDT in the 2lSS channel

variable	description
N_{jets}	Number of central jets with $p_T > 25$ GeV
$\Delta R(\ell_0, \text{jet})$	Angular distance between leading lepton and its closest jet
$\Delta R(\ell_1, \text{jet})$	Angular distance between sub-leading lepton and its closest jet
$M(\ell_0, \ell_1)$	Invariant mass of leading lepton and sub-leading lepton
LD	Linear discriminant defined as: $0.6 * E_T^{miss} + 0.4 * H_{\text{T}}^{\text{jet}}$[7]
$p_T(\text{jet}_0)$	Transverse momentum of the leading jet
$p_T(\text{jet}_1)$	Transverse momentum of the sub-leading jet
$\Delta R_{jets}^{\text{avg}}$	Average ΔR between jets
$\text{Max}(\lvert\eta_l\rvert)$	Pseudo-rapidity difference between the leading and subleading leptons ($\lvert\eta_{\ell_0}\rvert$ and $\lvert\eta_{\ell_1}\rvert$)
$p_T(\ell_1)$	Transverse momentum of the subleading lepton
$\eta(\ell_0)$	Pseudo-rapidity of the leading lepton
M(lep, MET)	Invariant mass of leptons and missing transverse energy
$M_T^{(\ell 0, MET)}$	Transverse mass of the leading lepton and missing transverse energy
$M_T^{(\ell 1, MET)}$	Transverse mass of the sub-leading lepton and missing transverse energy
$\eta(\text{jet}_0)$	Pseudo-rapidity of the leading jet
$\eta(\text{jet}_1)$	Pseudo-rapidity of the sub-leading jet
$H_{\text{T}}^{\text{jet}}$	Scalar sum of the transverse momenta of the jets
$\Delta R(\ell_0, \ell_1)$	Angular distance between the two same-sign leptons
M_{b0}	Invariant mass of the leading b-jet
M_{b1}	Invariant mass of the sub-leading b-jet

Definitions of the MVA-based CRs and SRs used in the 2lSS channel

Regions	Selections
Signal Regions	
$t\bar{t}H$ ++	$2\ell SS + 0\tau_{\text{had}}$ pre-MVA selection (cf. Table 7) total charge > 0 $\text{Max}(\text{BDT}_{t\bar{t}H}, \text{BDT}_{tH}, \text{BDT}_{t\bar{t}W}, \text{BDT}_{\text{Other}}) = \text{BDT}_{t\bar{t}H}$
$t\bar{t}H$ −−	$2\ell SS + 0\tau_{\text{had}}$ pre-MVA selection total charge < 0 $\text{Max}(\text{BDT}_{t\bar{t}H}, \text{BDT}_{tH}, \text{BDT}_{t\bar{t}W}, \text{BDT}_{\text{Other}}) = \text{BDT}_{t\bar{t}H}$
tH	$2\ell SS + 0\tau_{\text{had}}$ pre-MVA selection $\text{Max}(\text{BDT}_{t\bar{t}H}, \text{BDT}_{tH}, \text{BDT}_{t\bar{t}W}, \text{BDT}_{\text{Other}}) = \text{BDT}_{tH}$
Control Regions	
$t\bar{t}W$ ++	$2\ell SS + 0\tau_{\text{had}}$ pre-MVA selection total charge > 0 $\text{Max}(\text{BDT}_{t\bar{t}H}, \text{BDT}_{tH}, \text{BDT}_{t\bar{t}W}, \text{BDT}_{\text{Other}}) = \text{BDT}_{t\bar{t}W}$
$t\bar{t}W$ −−	$2\ell SS + 0\tau_{\text{had}}$ pre-MVA selection total charge < 0 $\text{Max}(\text{BDT}_{t\bar{t}H}, \text{BDT}_{tH}, \text{BDT}_{t\bar{t}W}, \text{BDT}_{\text{Other}}) = \text{BDT}_{t\bar{t}W}$
Other	$2\ell SS + 0\tau_{\text{had}}$ pre-MVA selection $\text{Max}(\text{BDT}_{t\bar{t}H}, \text{BDT}_{tH}, \text{BDT}_{t\bar{t}W}, \text{BDT}_{\text{Other}}) = \text{BDT}_{\text{Other}}$

Definitions of the 3l MVA regions

Regions	Selections
$t\bar{t}H$ SR	$t\bar{t}H > 0.2.$
tH SR	$tHjb > 0.25, t\bar{t}H < 0.2.$
$t\bar{t}W$ CR	$t\bar{t}W > 0.3, t\bar{t}H < 0.2, tHjb < 0.25.$
$t\bar{t}Z$ CR	$t\bar{t}Z > 0.45, t\bar{t}H < 0.2, tHjb < 0.25, t\bar{t}W < 0.3.$
VV CR	$VV > 0.65, t\bar{t}H < 0.2, tHjb < 0.25, t\bar{t}W < 0.3, t\bar{t}Z < 0.45.$
$t\bar{t}$ Region	$t\bar{t} > 0.25, t\bar{t}H < 0.2, tHjb < 0.25, t\bar{t}W < 0.3, t\bar{t}Z < 0.45, VV < 0.65.$
Other Region	$t\bar{t} < 0.25, t\bar{t}H < 0.2, tHjb < 0.25, t\bar{t}W < 0.3, t\bar{t}Z < 0.45, VV < 0.65.$

Event selection summary in the background control regions

Control regions for:	Diboson	$t\bar{t}Z$	Conversions	HF non-prompt
N_{jets}	2 or 3	≥ 4	≥ 0	≥ 2
$N_{b-\text{jets}}$	$1\ b^{85\%}$		$0\ b^{85\%}$	$1\ b^{85\%}$
Lepton requirement	3ℓ		$\mu\mu e^{*}$	2ℓSS
Lepton definition	(L, M, M)			(T, M_{ex}) —— (M_{ex}, T) —— $(M_{\text{ex}}, M_{\text{ex}})$
Lepton p_{T} [GeV]	(10, 15, 15)			(15, 15)
$\vert m^{\text{SF}}_{\ell^+\ell^-} - m_Z\vert$ [GeV]	< 10		> 10	–
$\vert m_{\ell\ell\ell} - m_Z\vert$ [GeV]	¿10		< 10	–
$m_T(\ell_0, E_{\text{T}}^{\text{miss}})$ [GeV]	–			< 250 for TM_{ex} and $M_{\text{ex}}T$ pairs
τ_{had} candidates (Medium)	0			0
Region split	–	–	internal / material	subleading $e/\mu \times (TM_{\text{ex}}, M_{\text{ex}}T, M_{\text{ex}}M_{\text{ex}})$
Region naming	3ℓVV	3ℓttZ	3ℓIntC	$2\ell\text{tt(e)}_{TM_{\text{ex}}}$, $2\ell\text{tt(e)}_{M_{\text{ex}}T}$, $2\ell\text{tt(e)}_{M_{\text{ex}}M_{\text{ex}}}$
			3ℓMatC	$2\ell\text{tt}(\mu)_{TM_{\text{ex}}}$, $2\ell\text{tt}(\mu)_{M_{\text{ex}}T}$, $2\ell\text{tt}(\mu)_{M_{\text{ex}}M_{\text{ex}}}$