Search for a Vector Like Quark T' decaying into quark top and Higgs boson in a dileptonic same sign final state with the CMS experiment for Run 2 at LHC

Top LHC France - 9 April 2024

Benjamin Blancon[1] on behalf of the CMS collaboration

[1]IP2I, Lyon, France

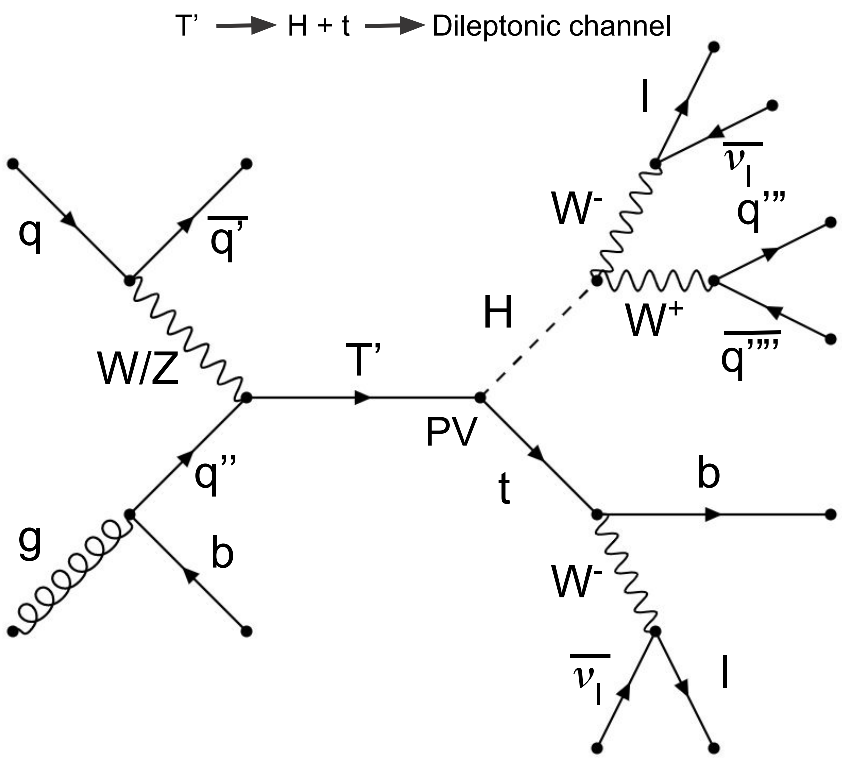
Motivation

- Vector-Like Quark (VLQ) T':
  - Present in many BSM models (Little Higgs models, Extra dimension models...).
  - Electroweak and strong interactions with SM particles.
  - Interaction with the Higgs boson solves the hierarchy problem.
  - Mass $M_{T'}$ predicted in many BSM models to be below or around 1 TeV for a Higgs boson mass $m_H \approx 125$ GeV (13 TeV in the center of mass of the LHC for 2016-2018).

- No evidence for an excess using full Run 2 data in the T' ⟶ tH all-hadronic resolved final state (see Stéphanie's presentation) and diphoton final state.

⟶ Needs to be confirmed in other decays!

Signal process: final state



- Final state:
  - Two leptons same sign (SS).
  - Three jets with one b-jet.
  - Two neutrinos.

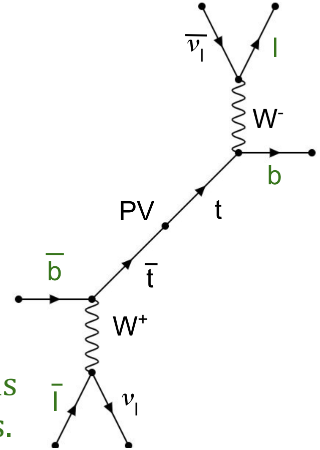
- Three channels:
  - μμ.
  - eμ.
  - ee.

- Nominal masses $M_{T'}$ range: [600 - 1200 GeV].

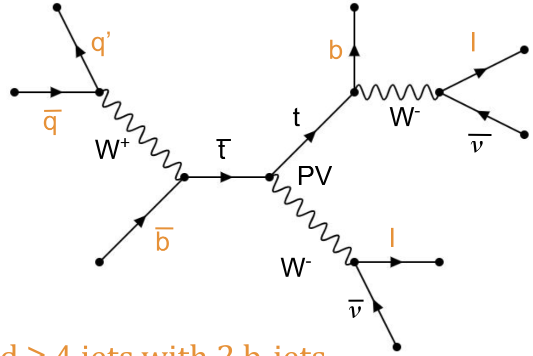
- Analysis performed with the full Run 2 (2016-2018) data.

Analysis strategy

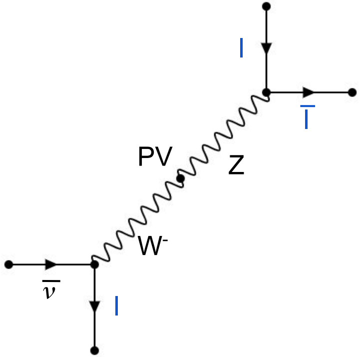
- Searching for a resonance in the transverse mass $M_{T,T'}$.
- Three Control regions implemented.
- Background estimated with data-driven technique.
- Combined fit of Signal region and Control regions.
- Optimized for $M_{T'} = 700$ GeV.



$t\bar{t}$ dileptonic ($t\bar{t}$): 1 or 2 leptons opposite sign (OS) and 2 b-jets.



$t\bar{t}W$ ($t\bar{t}X$): 2 leptons SS and ⩾ 4 jets with 2 b-jets.



WZ (Diboson): ⩾ 2 leptons.

Object identification

Muon	Pt > 30 GeV, $\|\eta\| < 2.4$, tight ID, optimized selection on isolation, distance to Primary Vertex (PV), 3D Impact parameter significance (SIP_{3D}) and charge misreconstruction veto.
Electron	Pt > 40 GeV, $\|\eta\| < 2.5$, tight ID, optimized selection on isolation, distance to PV, SIP_{3D}, number of hits, photon conversion veto and charge misreconstruction veto.
Jet	Pt > 30 GeV, $\|\eta\| < 2.4$, tight ID and removed overlap between the jets and the leptons in a cone of $\Delta R < 0.4$.
B-jet	Pt > 30 GeV, $\|\eta\| < 2.4$, tight ID and medium DeepJet WP. Identical overlap condition as the jets.

- Recommended CMS corrections applied (Pileup and b-tag weights, JEC, muon and electron efficiencies).

Signal selection

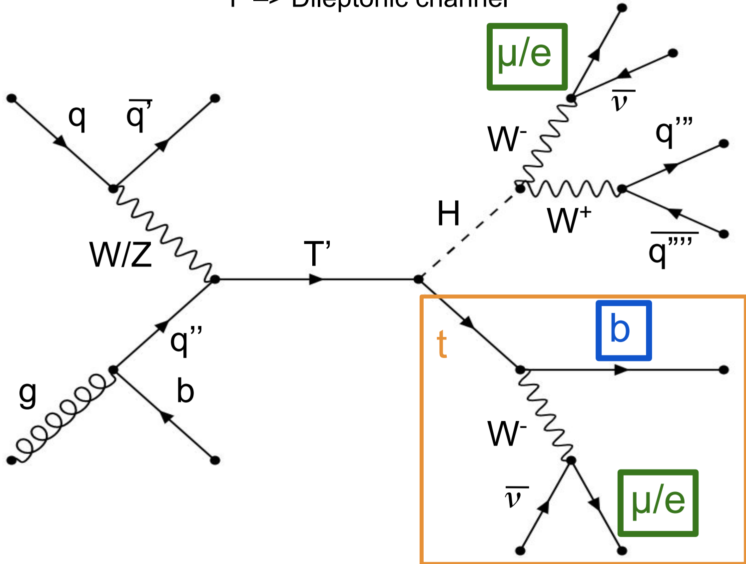
- Selection criteria:
  - The two leptons must be back-to-back.
  - The T' has a large mass so we expect the 2 leptons and the b-jet to have high Pt.
  - The top quark must have a non-hadronic decay.

- Falling $M_{T,T'}$ distribution needed to estimate the background.

→ Background efficiency in each $M_{T,T'}$ bin should be the same after each cut to keep the same shape!

→ Solution: cut optimization using quantiles of the background yield!

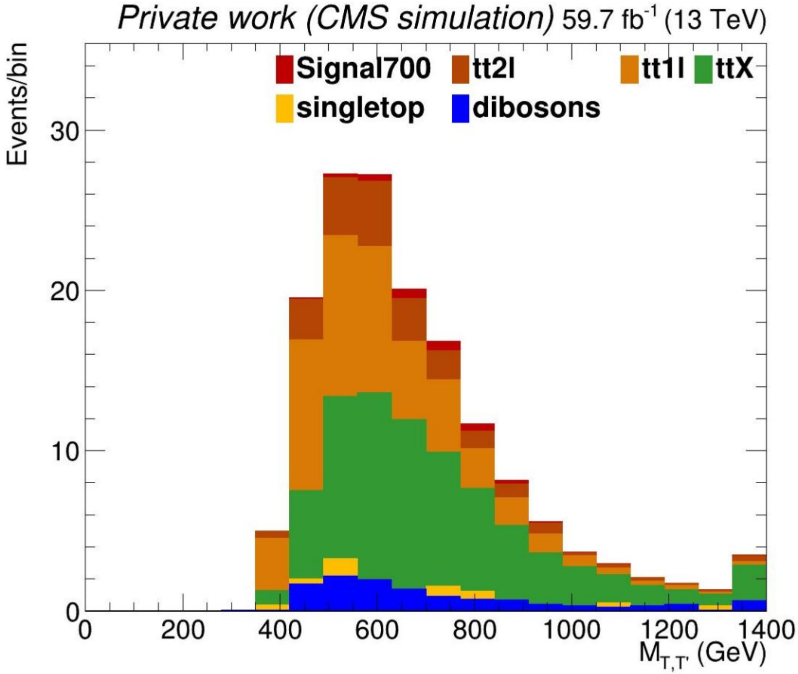
T' –> Dileptonic channel



Signal region (2018)

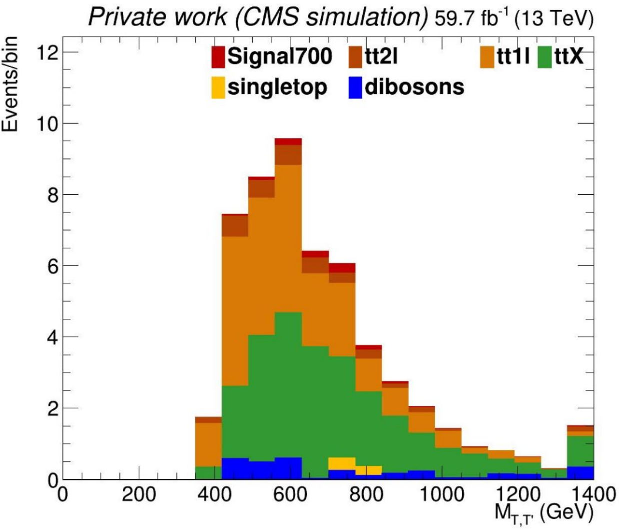
- All channels (μμ/eμ/ee) with 2018 simulation considered.

- Signal visible above the background (S/B = 1.83%).

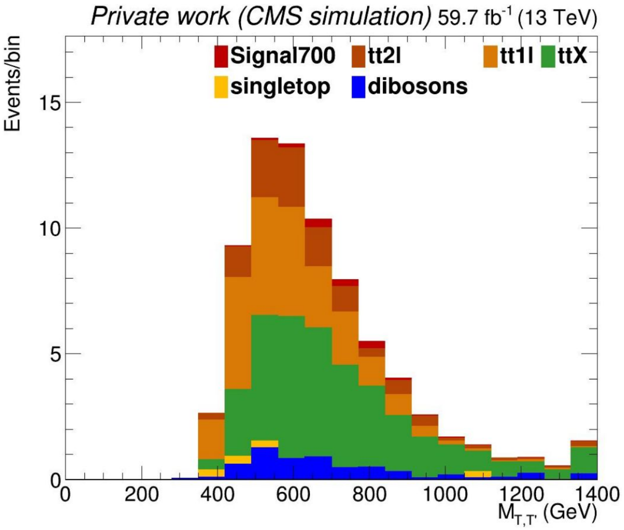


dibosons = WW + WW Deep Scattering + WWW + WH + ZH + WZ + ZZ
ttX = ttW + ttH + ttZ

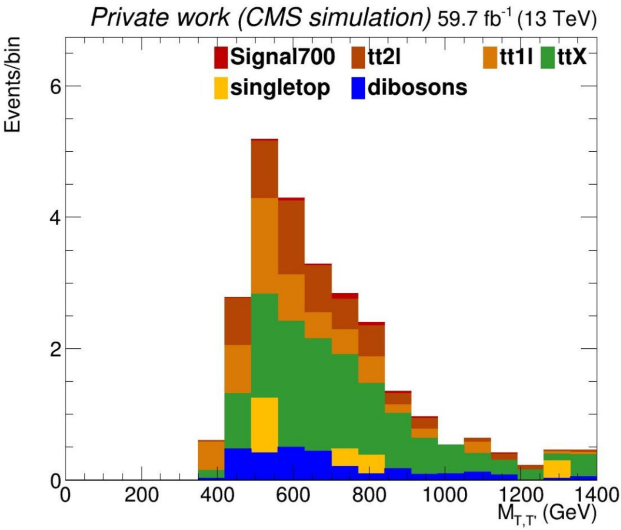
Signal region (2018) splitted by channel



μμ channel



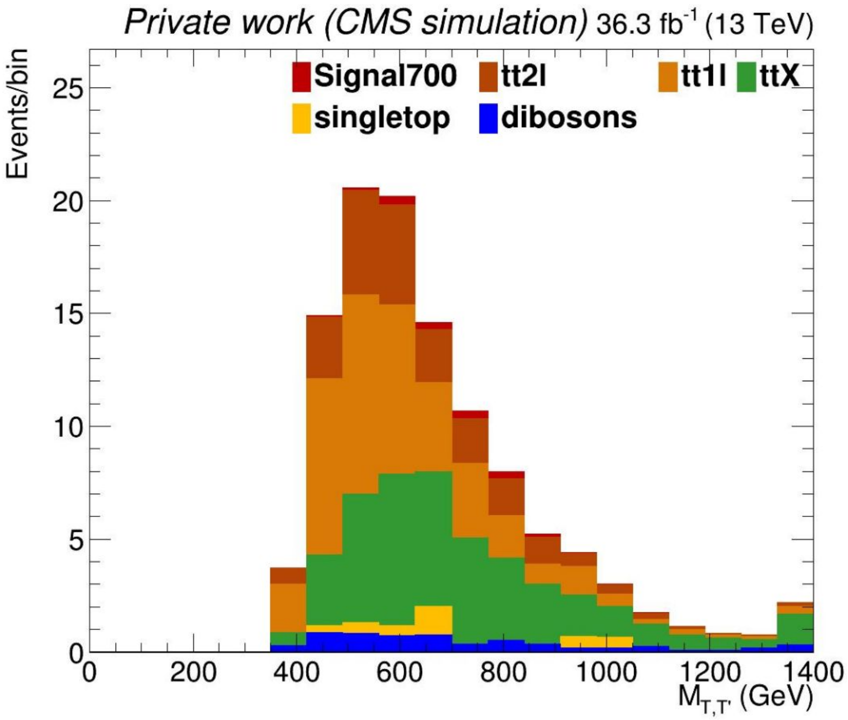
μe channel



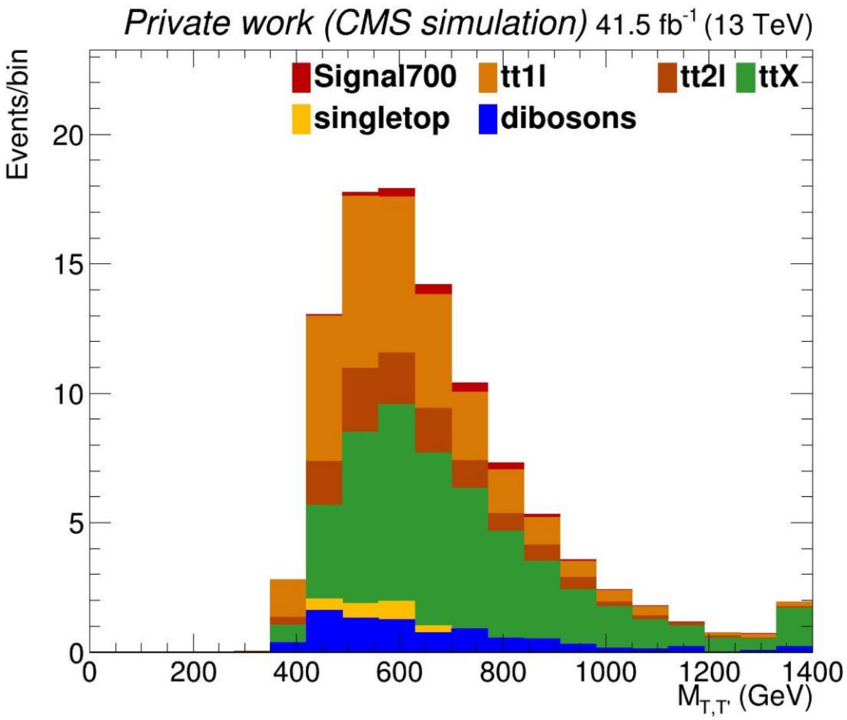
ee channel

- Higher signal over background discrimination for muonic channels (S/B = 2.06%, 1,87% and 1,28% for μμ, μe and ee channels respectively).

Signal region (2016 and 2017)



2016

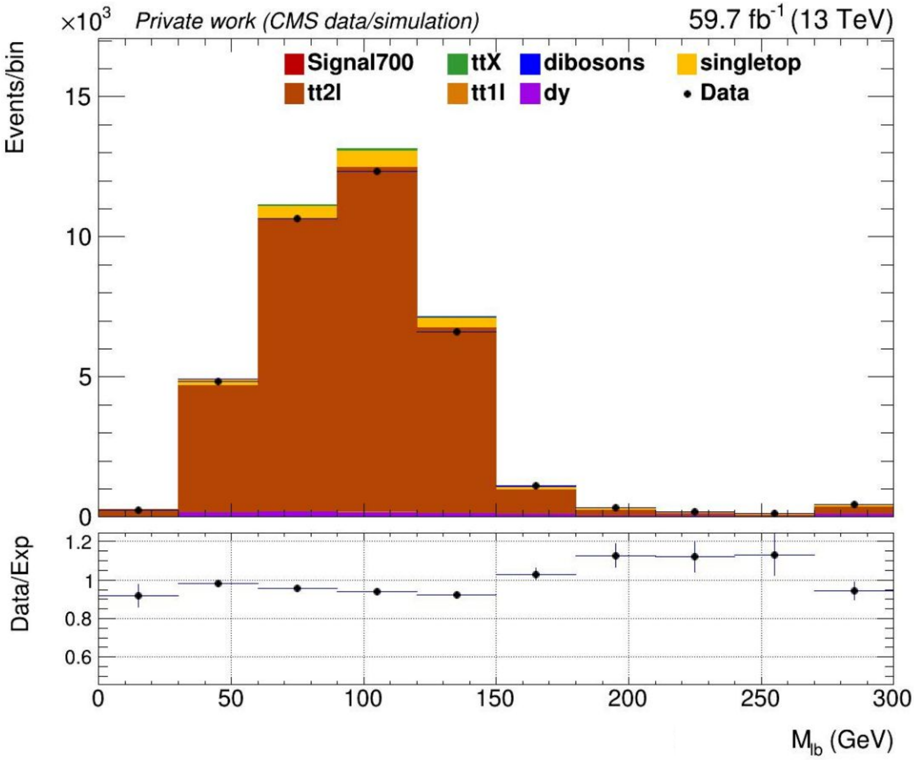


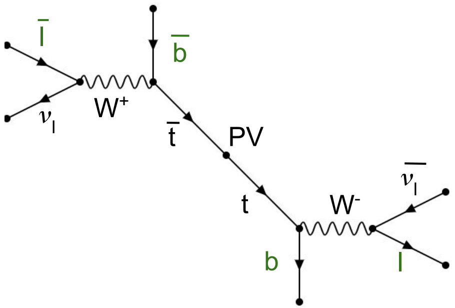
2017

- Same conclusion for 2016 (S/B = 1.57%) and 2017 (S/B = 1.84%).

Control regions - tt2l (2018)

- CR tt2l: two leptons with opposite sign.
- Discriminant variable in this Region: invariant mass of the lepton and the b-jet coming from the top M_{lb} (the neutrino is not included). Lepton and b-jet selected as the closest each other (correct identification in 80% of cases).
- Data/MC agreement within 1σ for all the studied variables.

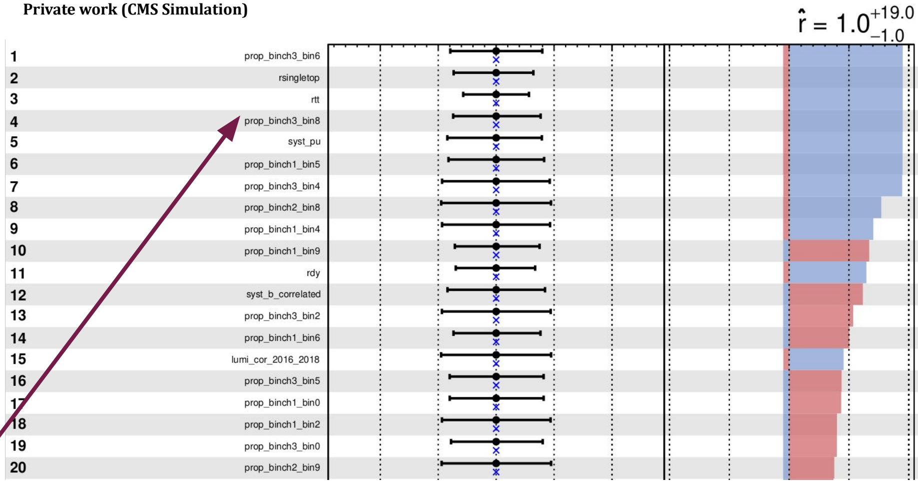




Nuisance pools - CR tt2l (2018)

- Impact of nuisance parameters:
  - All channels (μμ/eμ/ee) with 2018 simulation considered.
  - Signal injection test.
  - Systematic uncertainties: electron and muon efficiencies, b-tagging, rate on processes, luminosity, pileup and prefiring.
  - Statistical uncertainties implemented with the Barlow-Beeston-lite approach.

- Main constraints are single top and $t\bar{t}$ rate.

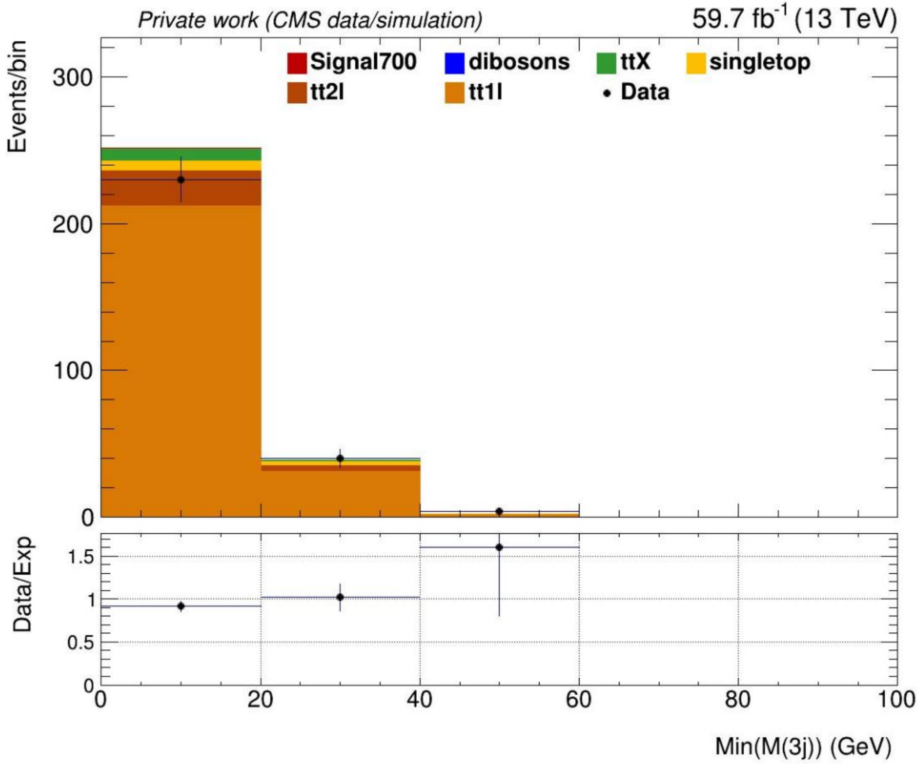


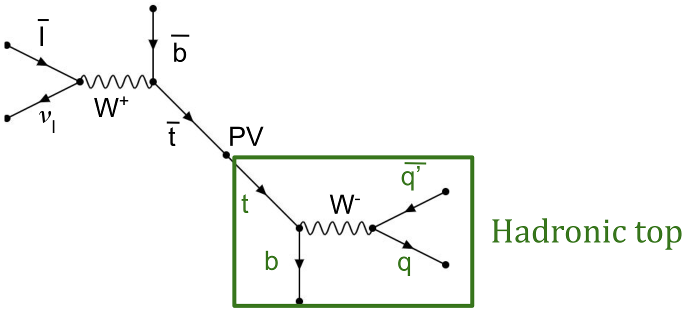
Control regions - tt1l (2018)

- CR tt1l: one prompt lepton and one jet faking lepton + revert hadronic top requirement wrt SR.

- Discriminant variable in this Region: invariant mass closest to the nominal top mass for the hadronic top among jet permutations Min(M(3j)).

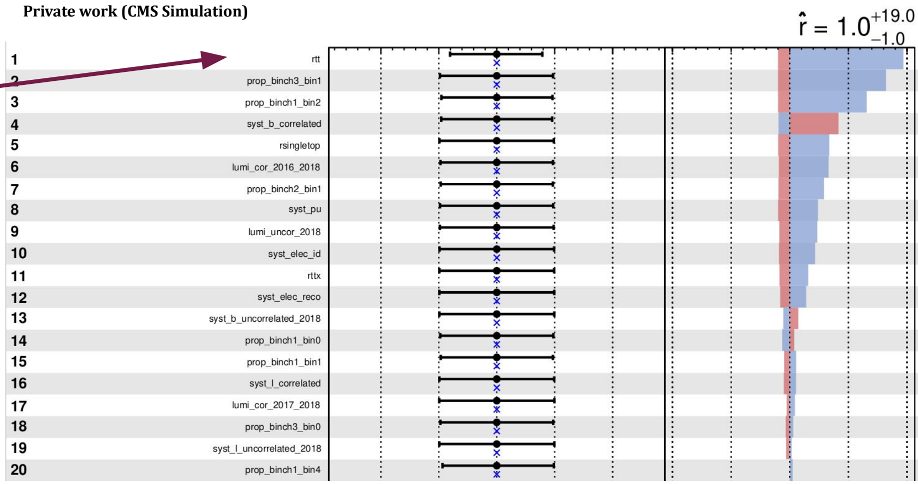
- Data/MC agreement within 1σ for all the studied variables but low statistics.





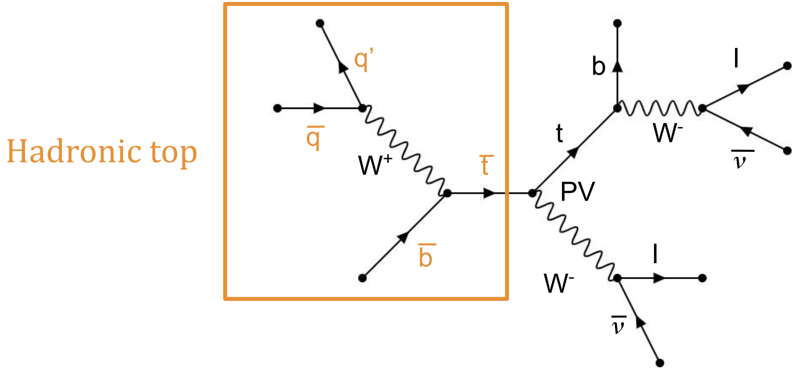
Nuisance pools - CR tt1l (2018)

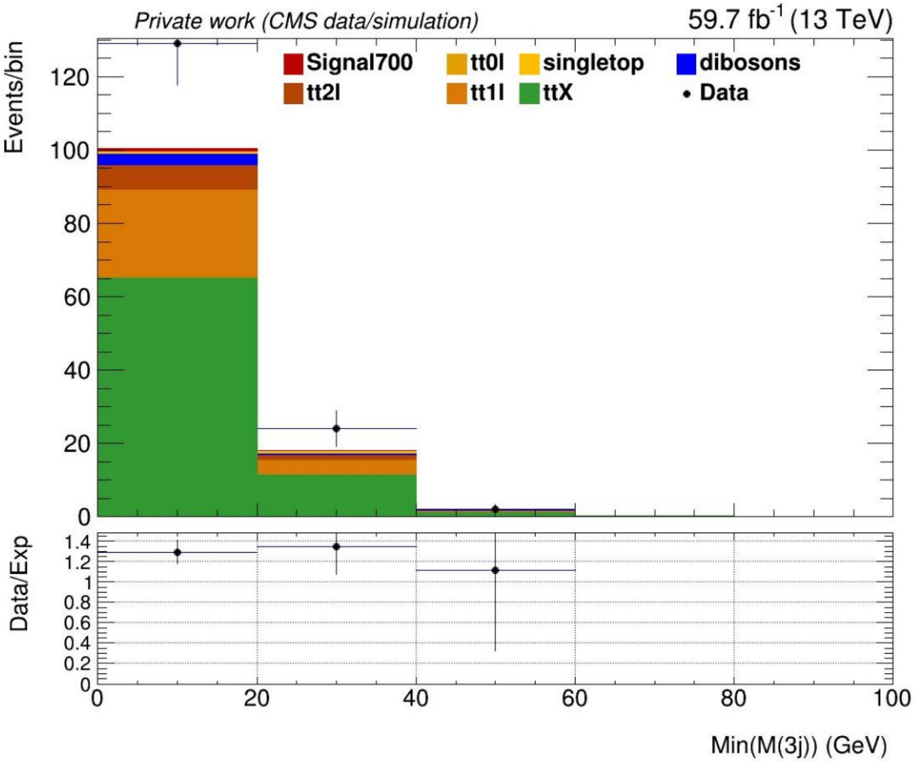
- Same procedure as previous CR.
- Main uncertainty is $t\bar{t}$ rate.



Control regions - ttX (2018)

- CR ttX: revert hadronic top requirement wrt SR.
- Discriminant variable in this Region: Min(M(3j)).
- MC underestimates data by 20% (maybe due to higher-order $t\bar{t}W$ corrections?).



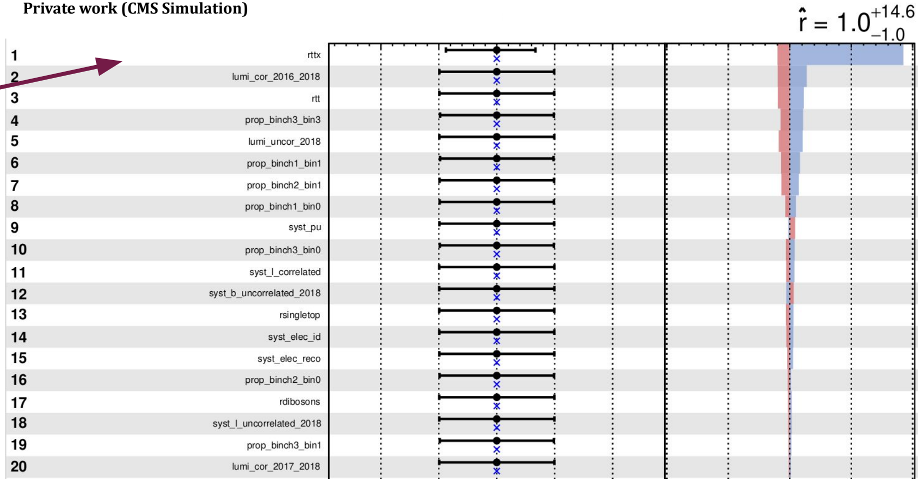


Nuisance pools - CR ttX (2018)

- Same procedure as previous CRs.
- Major uncertainty is $t\bar{t}X$ rate (by far).
- Strong anti-correlation between signal and $t\bar{t}X$ is observed.

→ New optimization needed.

Private work (CMS Simulation)



Correlation matrix of fit parameters

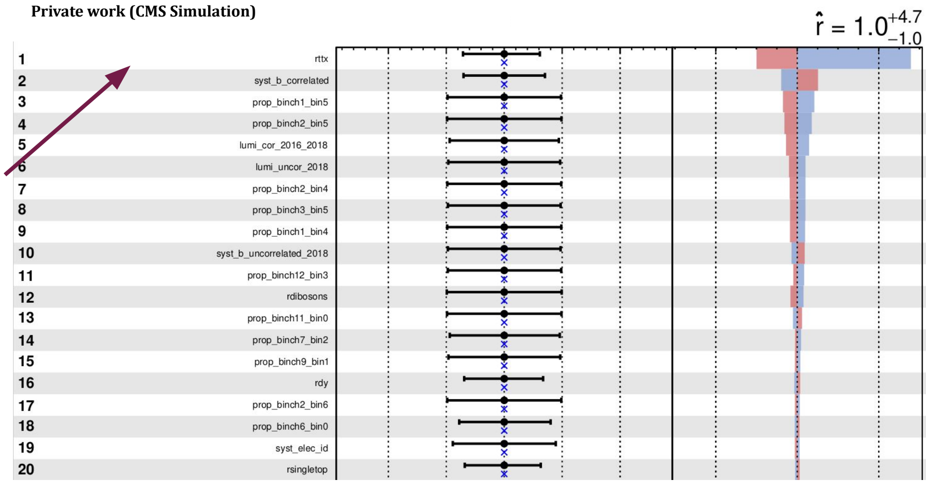
	r	rdibosons	rsingletop	rtt	rttx
r	1	-0.0166428	-0.0277781	-0.085852	-0.650929
rdibosons	-0.0166428	1	5.74209e-05	-0.000712093	-0.0068865
rsingletop	-0.0277781	5.74209e-05	1	-4.90835e-05	0.00159033
rtt	-0.085852	-0.000712093	-4.90835e-05	1	-0.0454508
rttx	-0.650929	-0.0068865	0.00159033	-0.0454508	1

Nuisance pools - All regions (2018)

- Impact plots with all channels and all regions for 2018 simulation considered.

- Major constraint and uncertainty still on $t\bar{t}X$ rate.

→ Most urgent issue to resolve.



Nuisance pools - All regions (Run 2)

- Impact plots with all channels and all regions for Run 2 simulation considered.

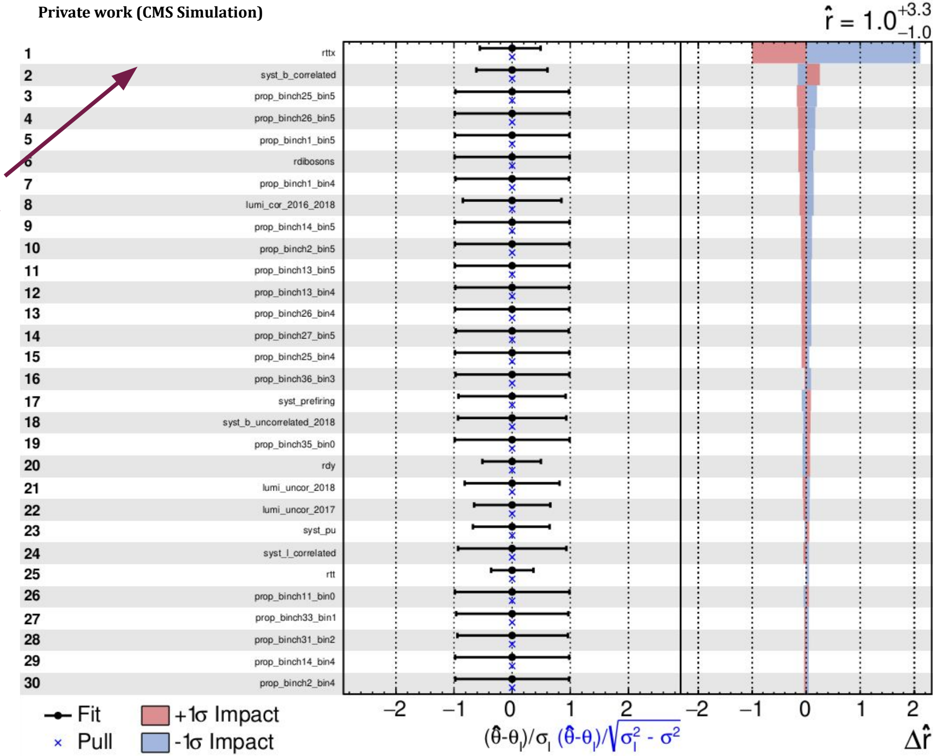
- Major constraint and uncertainty still on $t\bar{t}X$ rate.

→ Confirmation of the behavior in 2018.

- Several constraints to be looked at in more details.

- 95% upper limit on the cross-section:

$$\sigma = 587^{+241}_{-170} \text{ fb at } M_{T'} = 700 \text{ GeV.}$$



Summary

- Focus on T’ ⟶ top+H dileptonic SS final state using Run 2 samples.
- Analysis strategy settled: object identification, selection criteria and definition of the discriminant variables.
- Definition of the Control Regions, reasonable data/MC agreement.
- First impact of nuisances examined for full Run 2.
- Perspectives:
    - Apply remaining recommended CMS corrections to the simulated events.
    - Exclusion limits at 95% CL in the [600 GeV, 1200 GeV] range.


That's all Folks!

Back-up

VLQ production

- Several extensions of the Standard Model (SM) predict the existence of VLQs:
  - Spin ½ fermions.
  - Left-handed and right-handed components behave in the same way under the SM symmetry group.
  - Vector current couplings to the weak gauge bosons.
  - Non-Yukawa coupling mass-terms for VLQs are allowed.

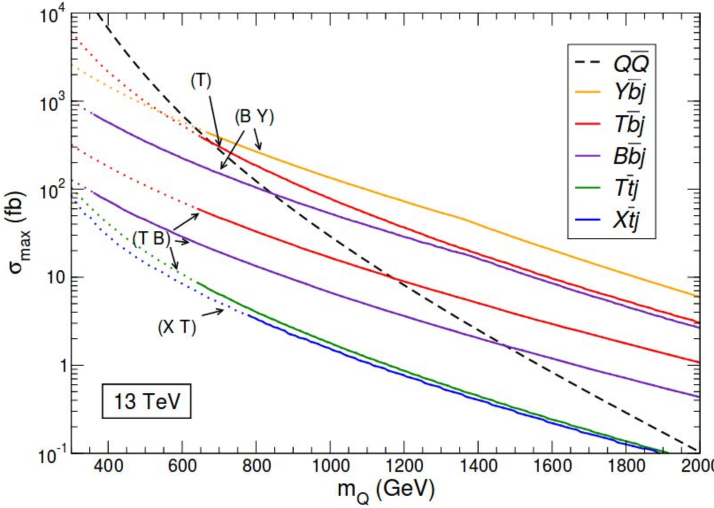
Type	Charge
X	+5/3
T	+2/3
B	-1/3
Y	-4/3

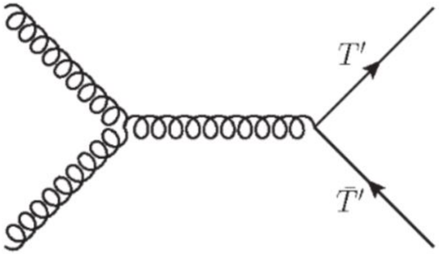
SU(2) Multiplets	
Singlets	T,B
Doublets	(T,B),(X,T),(B,Y)
Triplets	(X,T,B),(T,B,Y)

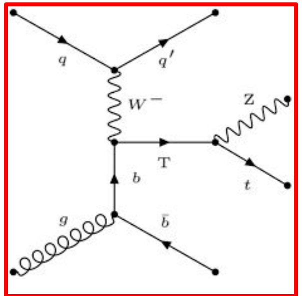
[Slides from Francesco]

VLQ production

- VLQs could be produced both singly and in pair:
  - Pair production:
    - Strong interaction processes.
    - Model independent cross section, suppressed for large VLQ mass.
  - Single production:
    - Electroweak processes.
    - Cross section depending on VLQ mass and coupling to SM particles.
    - Models foresee preferential mixing with 3rd generation SM quarks.

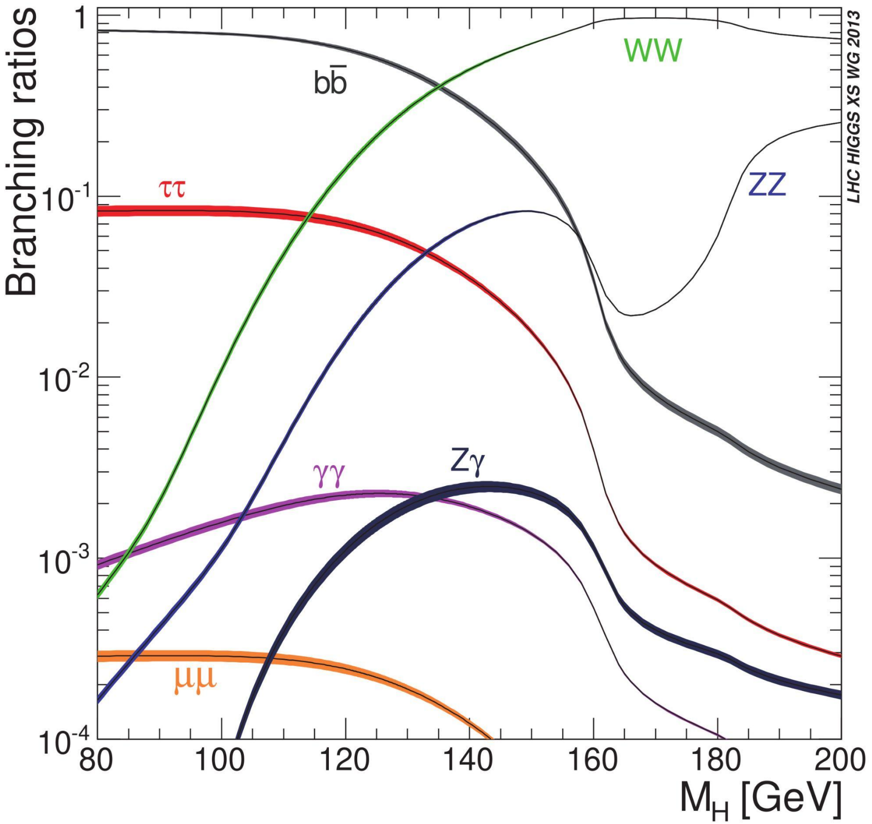






Type	Decay channel
X	tW
T	tZ,tH,tA,bW
B	bZ,bH,tW
Y	bW

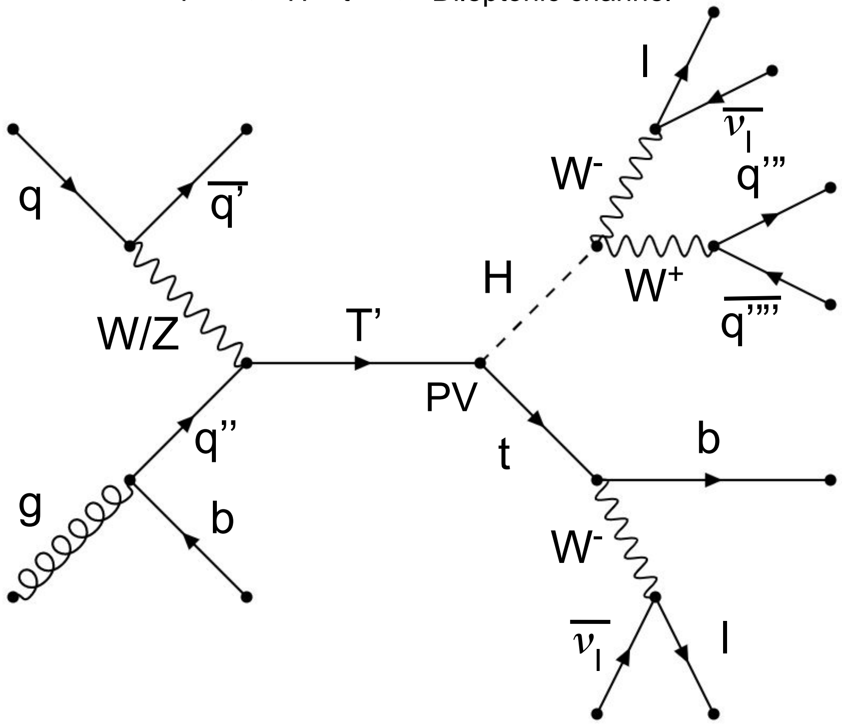
Higgs boson branching ratio decay



Reference

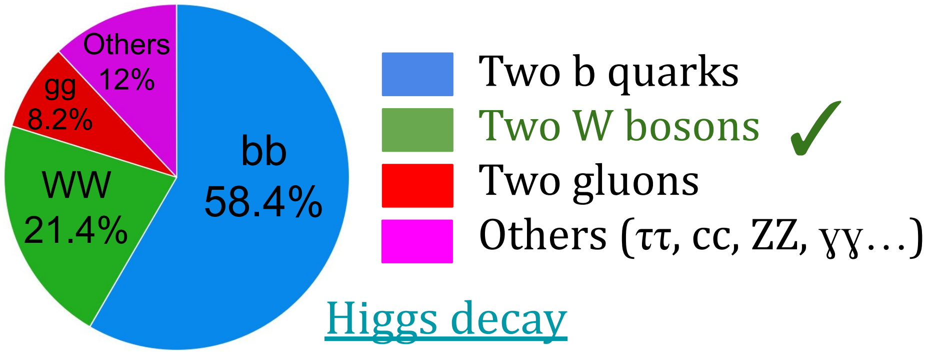
T’ production

T’ ⟶ H + t ⟶ Dileptonic channel



- Single T’ produced via electroweak process.

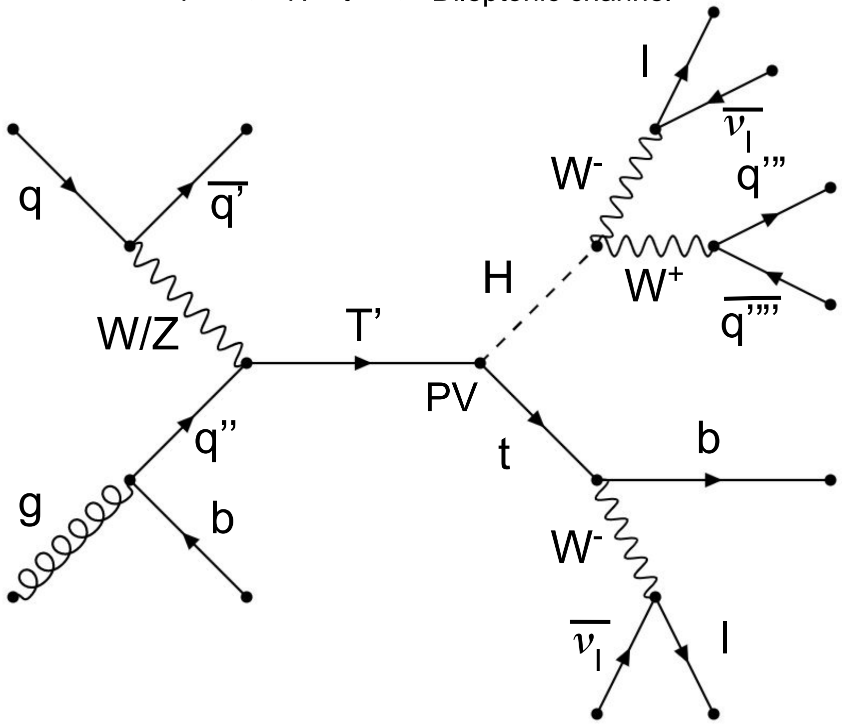
T’ decay



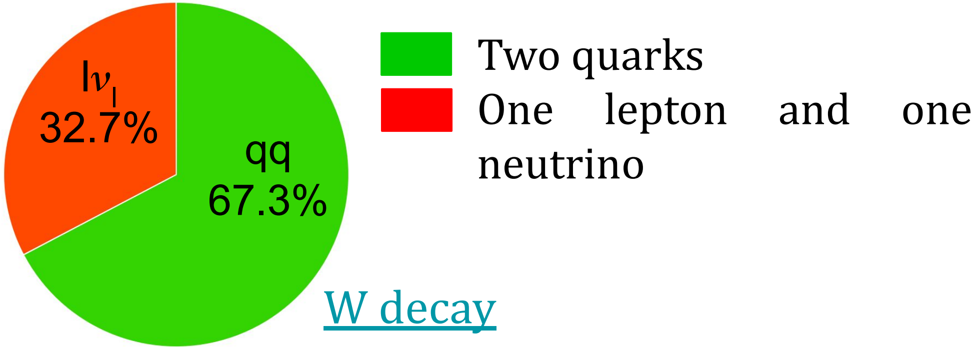
Higgs decay

T' production

T' ⟶ H + t ⟶ Dileptonic channel



- W leptonic or hadronic decay.



W decay

Corrections/Filters

- List of corrections already applied:
  - Generator weights.
  - Pileup weights.
  - B-tag weights.
  - Jet energy corrections (to be updated).
  - Muon SFs (Trigger/Reco/ID/Isolation/Rochester) (to be updated).
  - Electron SFs (Trigger/Reco/ID) (to be updated).

- List of corrections to apply in the future:
  - Jet energy resolution.

- Filters applied:
  - Flag_goodVertices
  - Flag_globalSuperTightHalo2016Filter
  - Flag_HBHENoiseFilter
  - Flag_HBHENoiseIsoFilter
  - Flag_EcalDeadCellTriggerPrimitiveFilter
  - Flag_BadPFMuonFilter
  - Flag_BadPFMuonDzFilter
  - Flag_hfNoisyHitsFilter
  - Flag_eeBadScFilter
  - Flag_ecalBadCalibFilter (2017/2018)

Discriminant variable

$M_{T,T'}$ is defined as follows.

$$M^2_{T,T'} = \left(\sum_{particles} E_T \right)^2 - \left(\sum_{particles} \overrightarrow{p_T} \right)^2$$
$$= \left(E_{T,lep,Higgs} + E_{T,\nu,Higgs} + E_{T,j1,Higgs} + E_{T,j2,Higgs} + E_{T,lep,top} + E_{T,\nu,top} + E_{T,b,top} \right)^2$$
$$- \left(\overrightarrow{p_{T,lep,Higgs}} + \overrightarrow{p_{T,\nu,Higgs}} + \overrightarrow{p_{T,j1,Higgs}} + \overrightarrow{p_{T,j2,Higgs}} + \overrightarrow{p_{T,lep,top}} + \overrightarrow{p_{T,\nu,top}} + \overrightarrow{p_{T,b,top}} \right)^2.$$

$$E_{T,\nu,Higgs} + E_{T,\nu,top} = \sqrt{\overrightarrow{p_{T,\nu,Higgs}}^2 + M^2_{T,\nu,Higgs}} + \sqrt{\overrightarrow{p_{T,\nu,top}}^2 + M^2_{T,\nu,top}}$$
$$= |\overrightarrow{p_{T,\nu,Higgs}}| + |\overrightarrow{p_{T,\nu,top}}|$$
$$= \sqrt{\left(|\overrightarrow{p_{T,\nu,Higgs}}| + |\overrightarrow{p_{T,\nu,top}}| \right)^2}$$
$$= \sqrt{|\overrightarrow{p_{T,\nu,Higgs}}|^2 + |\overrightarrow{p_{T,\nu,top}}|^2 + 2|\overrightarrow{p_{T,\nu,Higgs}}||\overrightarrow{p_{T,\nu,top}}|}.$$

$$\overrightarrow{p_{T,MET}}^2 = \left(\overrightarrow{p_{T,\nu,Higgs}} + \overrightarrow{p_{T,\nu,top}} \right)^2$$
$$= |\overrightarrow{p_{T,\nu,Higgs}}|^2 + |\overrightarrow{p_{T,\nu,top}}|^2 + 2|\overrightarrow{p_{T,\nu,Higgs}}||\overrightarrow{p_{T,\nu,top}}| \cos(\nu_H, \nu_t).$$

Hypothesis 1: the two neutrinos are anticolinear ($\cos(\nu_H, \nu_t) = -1$).

$$\Rightarrow E_{T,\nu,Higgs} + E_{T,\nu,top} = \sqrt{\overrightarrow{p_{T,MET}}^2 + 4 \times |\overrightarrow{p_{T,\nu,Higgs}}||\overrightarrow{p_{T,\nu,top}}|}.$$

$$\Rightarrow M_{T,T'} = \sqrt{\left(E_{T,lep,Higgs} + E_{T,j1,Higgs} + E_{T,j2,Higgs} + E_{T,lep,top} + E_{T,b,top} + \sqrt{\overrightarrow{p_{T,MET}}^2 + 4 \times |\overrightarrow{p_{T,\nu,Higgs}}||\overrightarrow{p_{T,\nu,top}}|} \right)^2 - \left(\overrightarrow{p_{T,ljjlb}} + \overrightarrow{p_{T,MET}} \right)^2}.$$

Discriminant variable

Hypothesis 2: $\left|\overrightarrow{p_{T,\nu,Higgs}}\right|$ is taken from the GEN information per mass point value. We have two cases.

- If $\left|\overrightarrow{p_{T,\nu,Higgs}}\right| > \left|\overrightarrow{p_{T,\nu,top}}\right|$,

$$\left|\overrightarrow{p_{T,MET}}\right| = \left|\overrightarrow{p_{T,\nu,Higgs}}\right| - \left|\overrightarrow{p_{T,\nu,top}}\right|.$$
$$\Leftrightarrow \left|\overrightarrow{p_{T,\nu,top}}\right| = \left|\overrightarrow{p_{T,\nu,Higgs}}\right| - \left|\overrightarrow{p_{T,MET}}\right|.$$

  As $\left|\overrightarrow{p_{T,\nu,top}}\right| > 0$, this case occurs when $\left|\overrightarrow{p_{T,\nu,Higgs}}\right| > \left|\overrightarrow{p_{T,MET}}\right|$.

- If $\left|\overrightarrow{p_{T,\nu,Higgs}}\right| < \left|\overrightarrow{p_{T,\nu,top}}\right|$,

$$\left|\overrightarrow{p_{T,MET}}\right| = \left|\overrightarrow{p_{T,\nu,top}}\right| - \left|\overrightarrow{p_{T,\nu,Higgs}}\right|.$$
$$\Leftrightarrow \left|\overrightarrow{p_{T,\nu,top}}\right| = \left|\overrightarrow{p_{T,\nu,Higgs}}\right| + \left|\overrightarrow{p_{T,MET}}\right|.$$

  As $\left|\overrightarrow{p_{T,\nu,top}}\right| > 0$, this case occurs when $\left|\overrightarrow{p_{T,\nu,Higgs}}\right| < \left|\overrightarrow{p_{T,MET}}\right|$.

$$\Rightarrow M_{T,T'} = \sqrt{\left(E_{T,lep,Higgs} + E_{T,j1,Higgs} + E_{T,j2,Higgs} + E_{T,lep,top} + E_{T,b,top} + \sqrt{\overrightarrow{p_{T,MET}}^2 + 4 \times \left|\overrightarrow{p_{T,\nu,Higgs}}\right| \times \left|\left|\overrightarrow{p_{T,\nu,Higgs}}\right| \pm \left|\overrightarrow{p_{T,MET}}\right|\right|}\right)^2 - \left(\overrightarrow{p_{T,ljjlb}} + \overrightarrow{p_{T,MET}}\right)^2}.$$

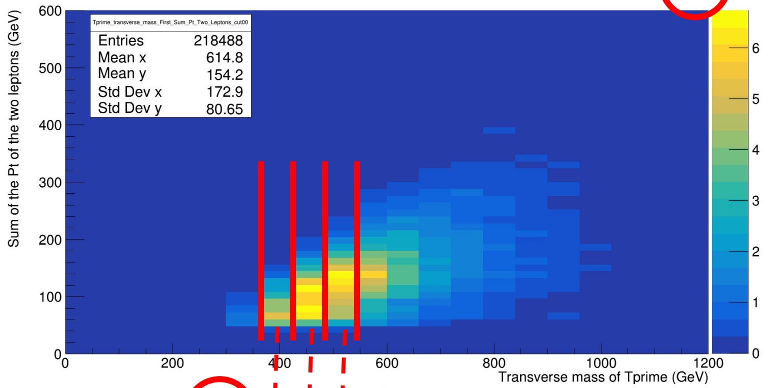
- Figure of merit: Significance defined as S = sqrt(2*lnQ) where $\ln Q = \sum_{\text{bins}} (s+b)\ast\ln(1+s/b) - s$.
- Run 2, All channels (μμ/eμ/ee):
  - Strategy 1: S = 0.467 σ.
  - Strategy 2: S = 0.307 σ.

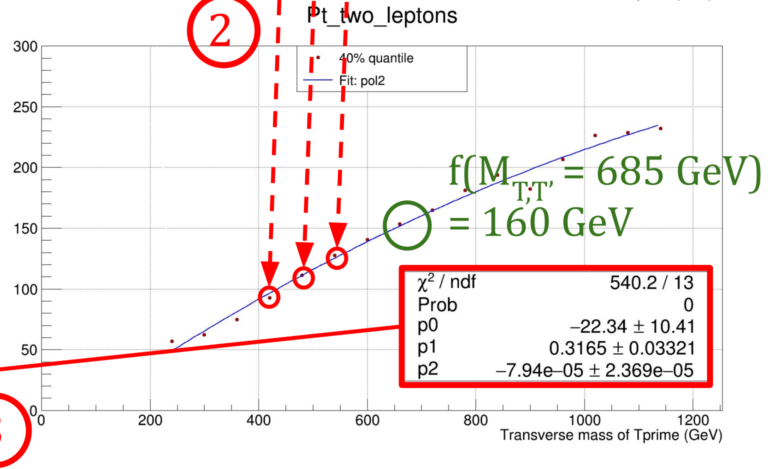
Signal selection - Optimization

Private work (CMS Simulation)

- First 'naive' cut to have the best S/B ratio for the target Cut variable (example: signal bump at $M_{T,T'}$ = 685 GeV for Cut variable > 160 GeV).
- Cut variable (y-axis) vs $M_{T,T'}$ (x-axis) is plotted in a 2D histogram for the background. 1
- Percentage of the background yield kept in each bin to plot the related 1D graph. The percentage is chosen so that the signal will be not modified, i.e. the 'naive' cut is still valid. 2
- Fit of the 1D graph.

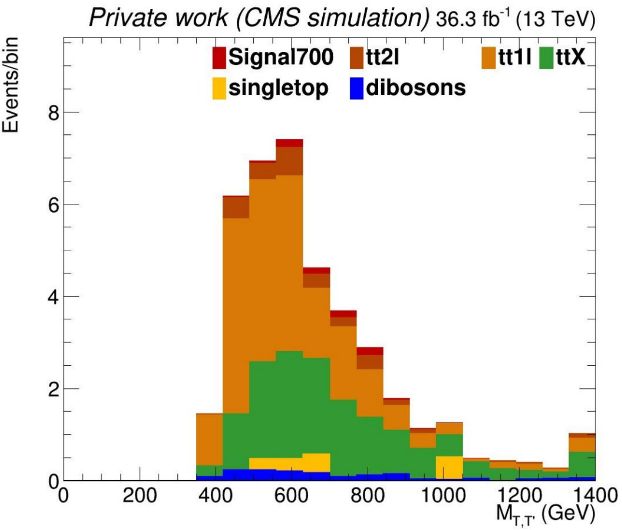
→ New Cut that preserves the background shape while keeping the signal strength! 3



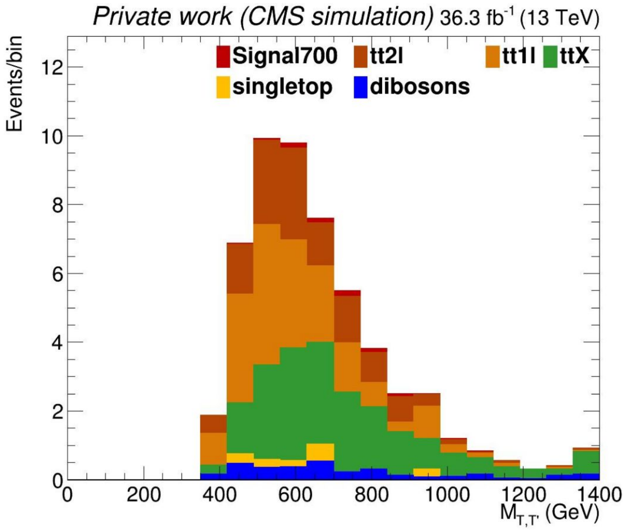


Cut variable > $f(M_{T,T'})$.

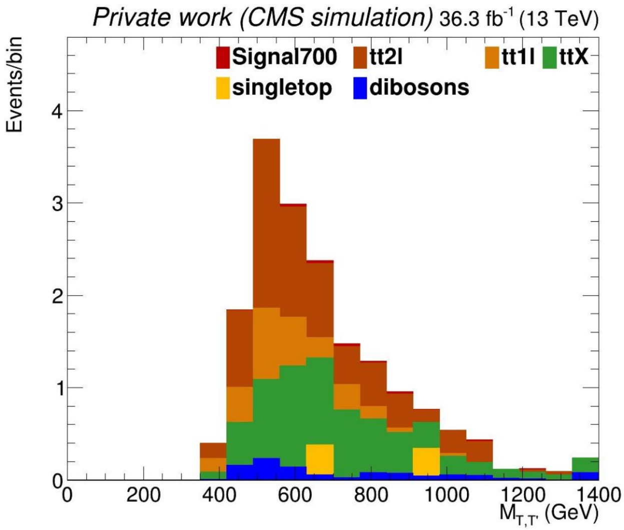
Signal region (2016) splitted by channel



μμ channel

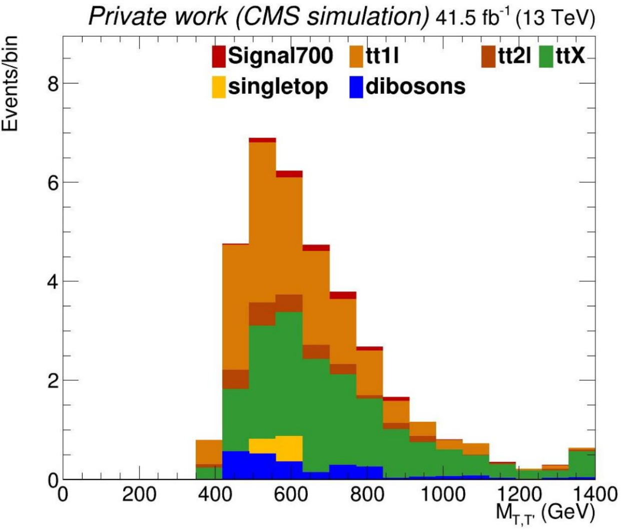


μe channel

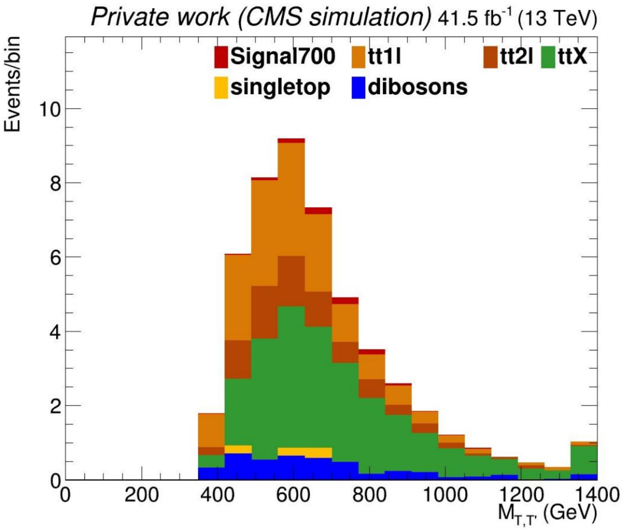


ee channel

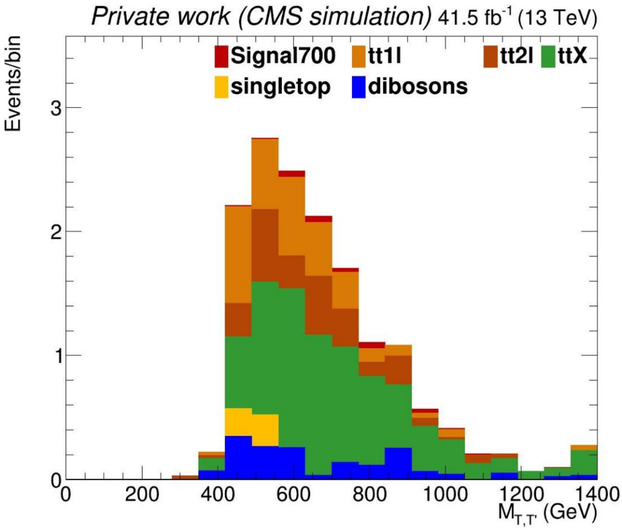
Signal region (2017) splitted by channel



μμ channel



μe channel



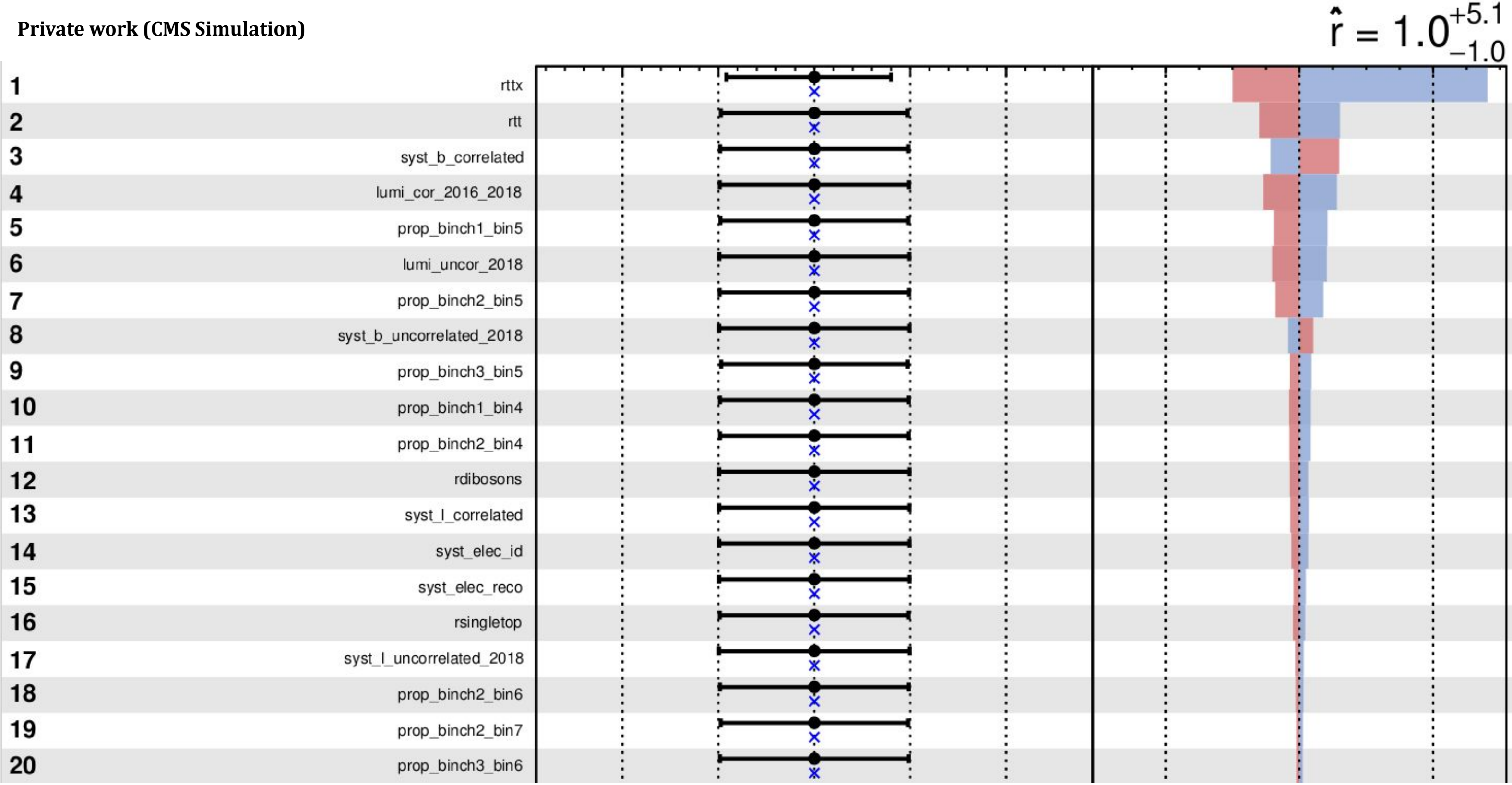
ee channel

Uncertainties

- Theoretical cross sections:
  - $t\bar{t}$: 6%.
  - Single top: 30%.
  - $t\bar{t}X$: 20%.
  - Dibosons: 10%.
  - Drell-Yan: 25%.
- Leptons:
  - Electron (Reco/ID) (systematics).
  - Muon (ID/Isolation) (systematics+statistics per year).
- Jets:
  - Heavy jets b-tagging (correlated/uncorrelated per year) (systematics).
  - Light jets b-tagging (correlated/uncorrelated per year) (systematics).
- Pileup (systematics).
- Prefiring (systematics).
- Luminosity (systematics):
  - Uncorrelated per year.
  - Correlated for 2016-2018 and 2017-2018.

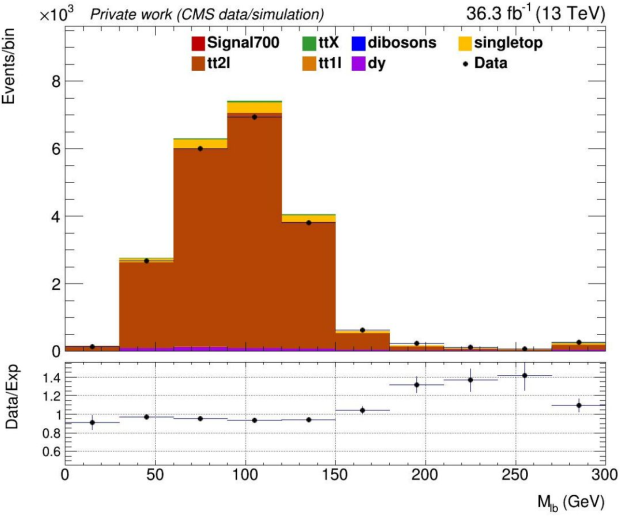
Nuisance pools - SR (2018)

Private work (CMS Simulation)

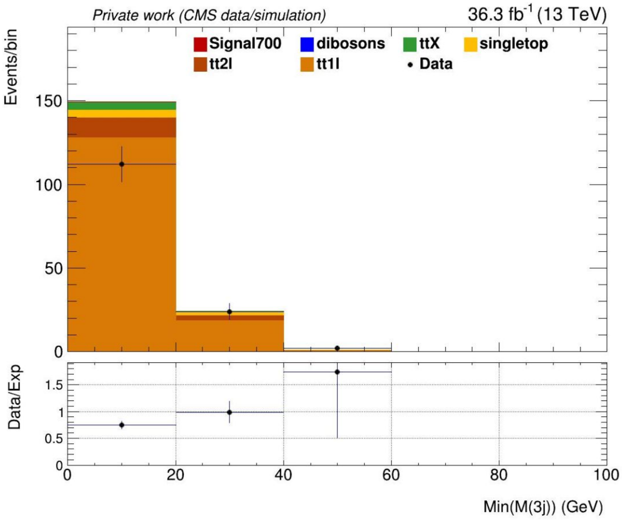


Benjamin Blancon
Top LHC France
04.09.24

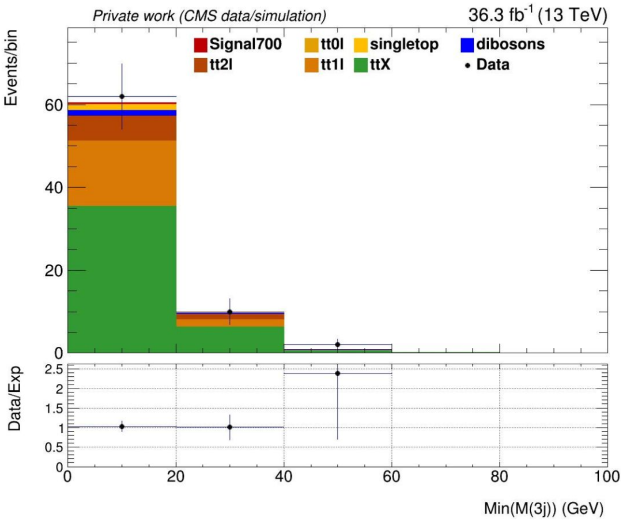
Control regions (2016)



CR tt2l



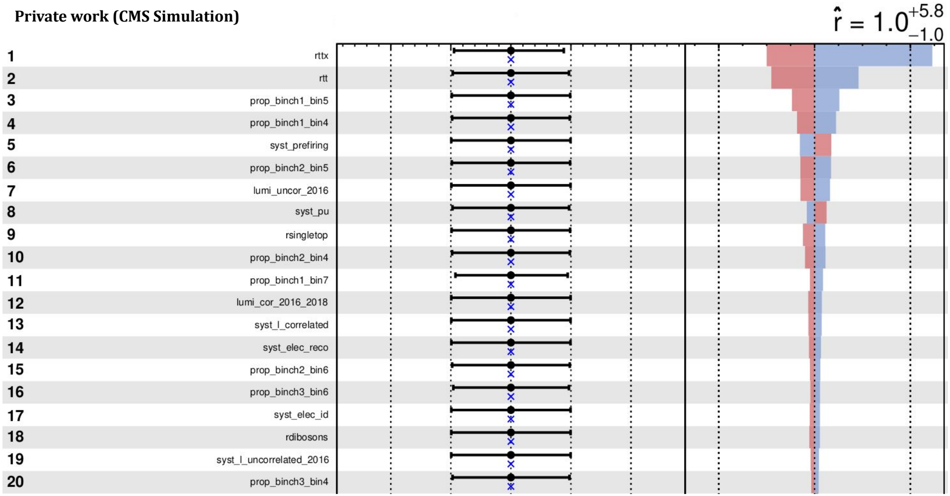
CR tt1l



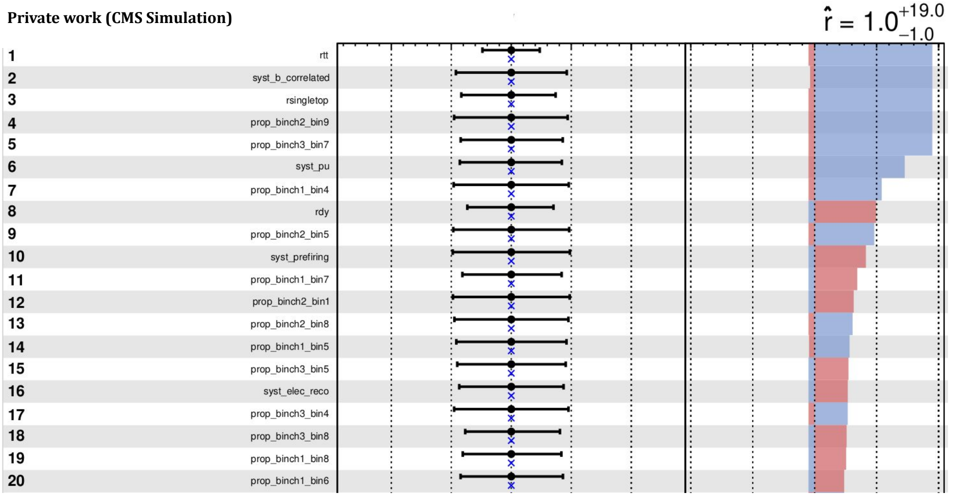
CR ttX

Benjamin Blancon
Top LHC France
04.09.24

Nuisance pools (2016)

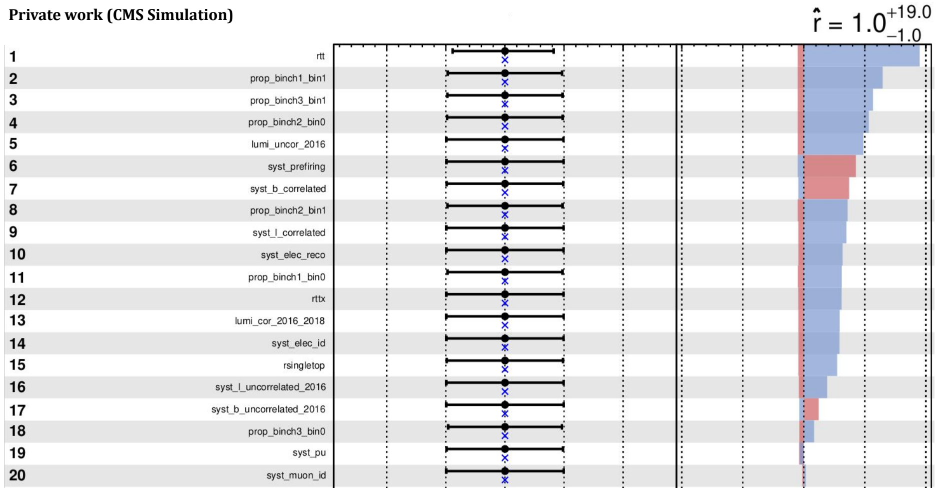


SR

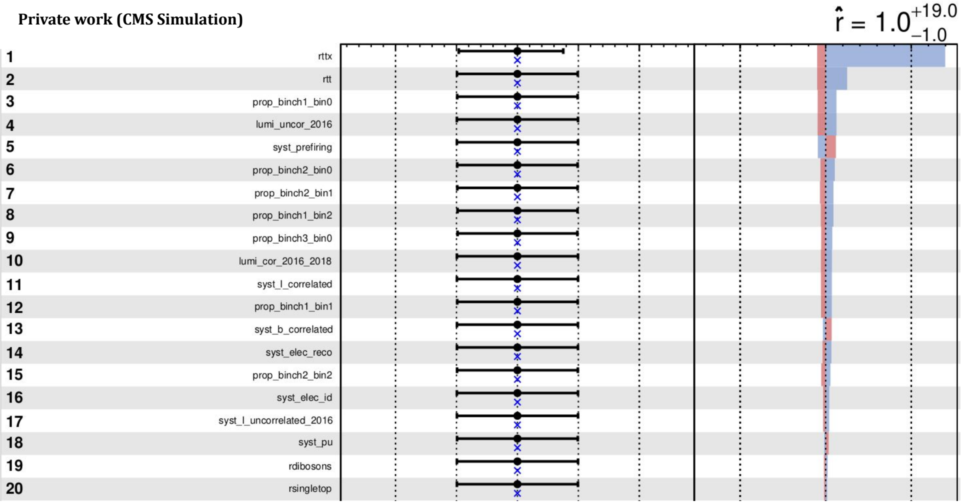


CR tt2l

Nuisance pools (2016)



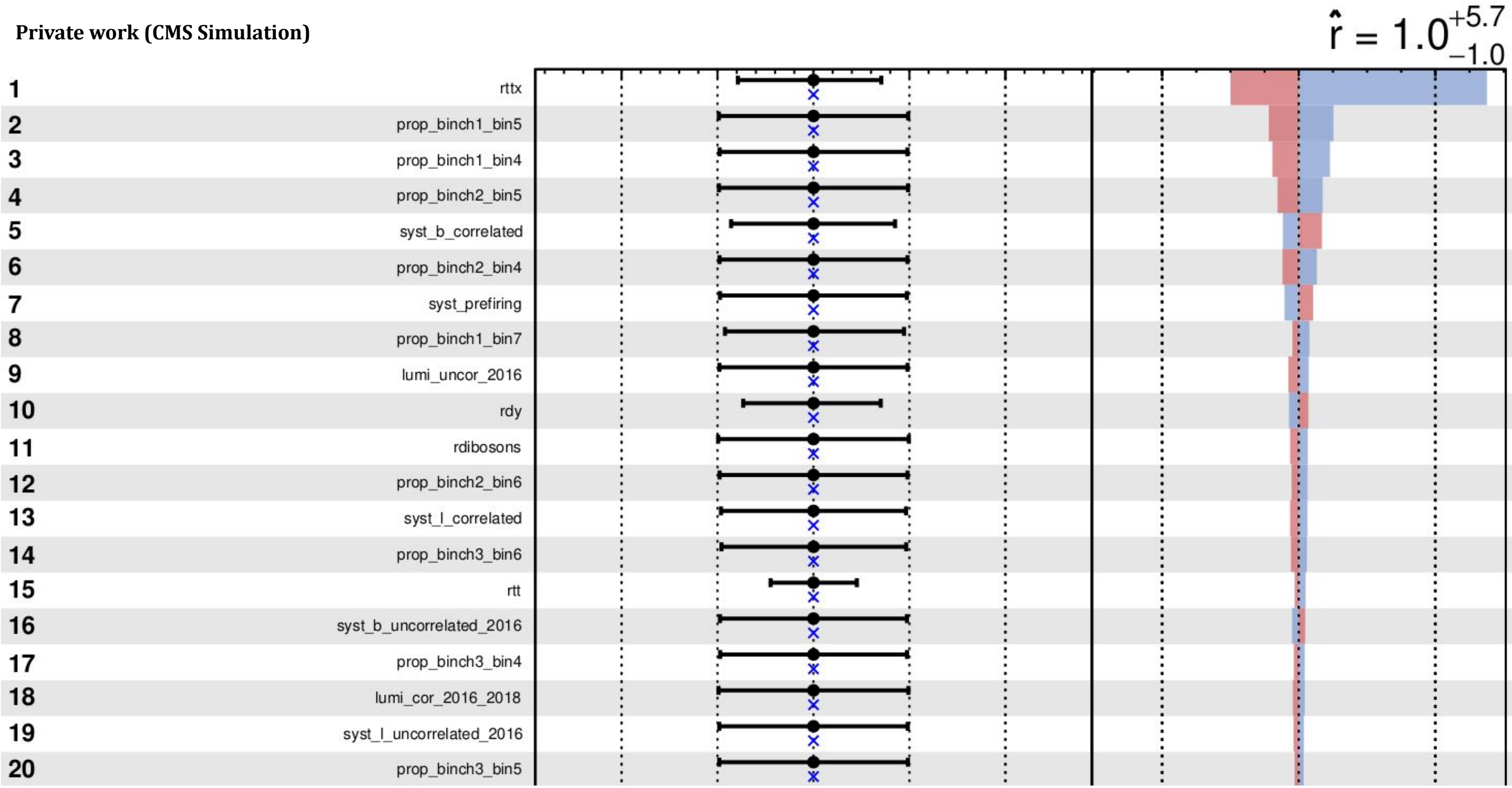
CR tt1l



CR ttX

Benjamin Blancon
Top LHC France
04.09.24

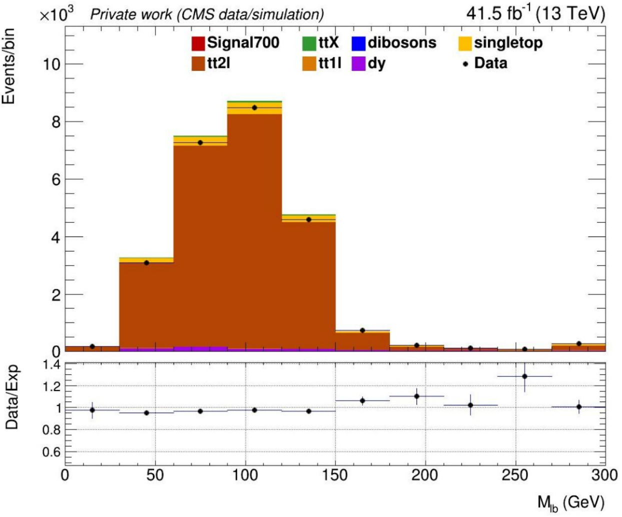
Nuisance pools (2016)



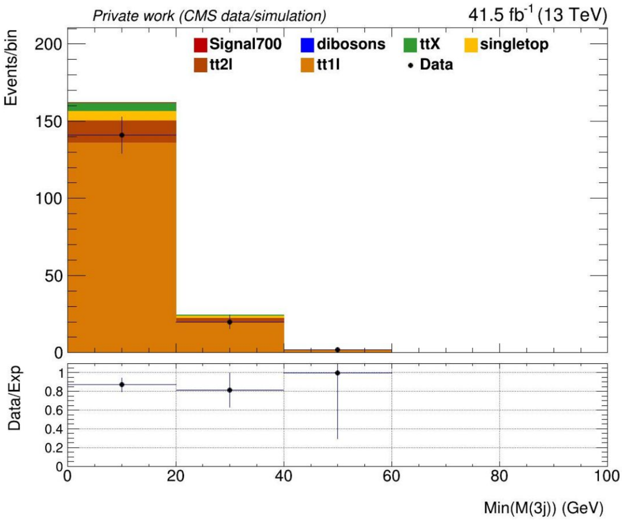
All regions

Benjamin Blancon
Top LHC France
04.09.24

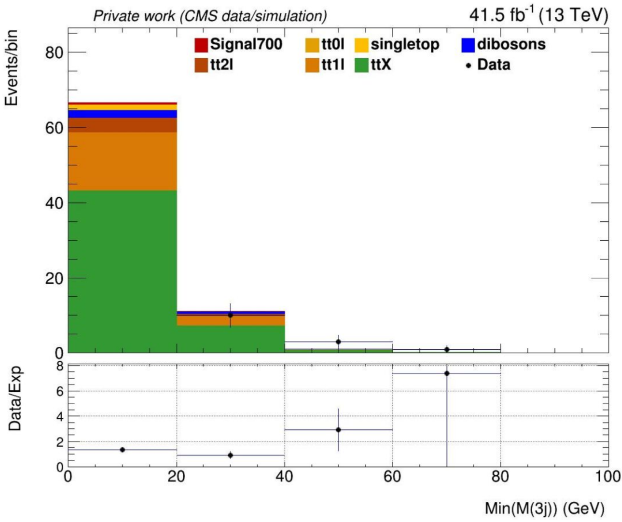
Control regions (2017)



CR tt2l



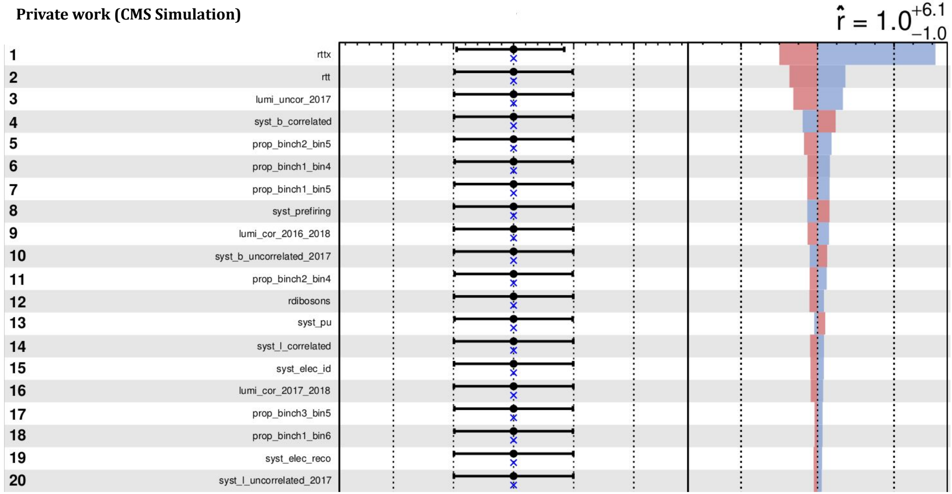
CR tt1l



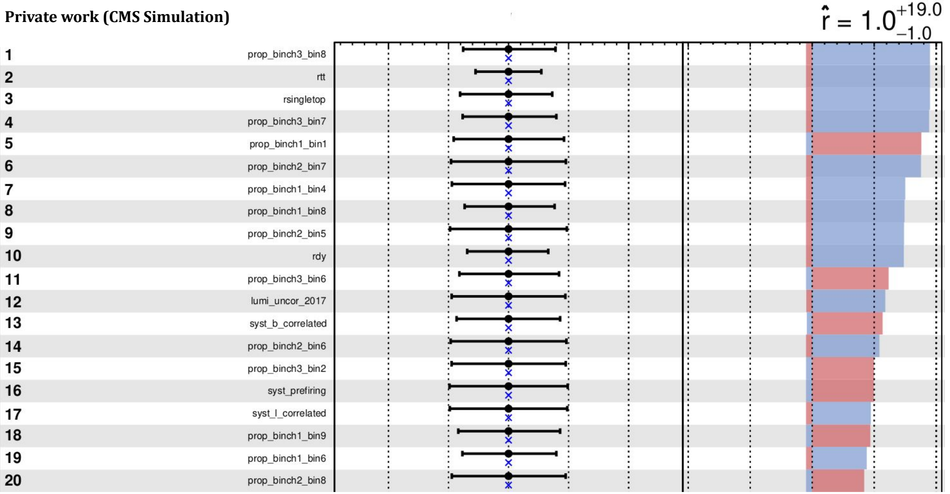
CR ttX

Benjamin Blancon
Top LHC France
04.09.24

Nuisance pools (2017)

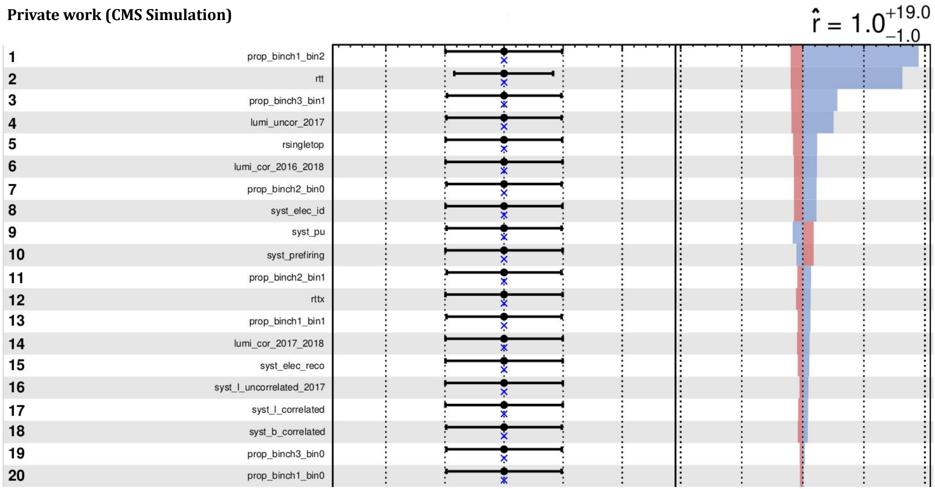


SR

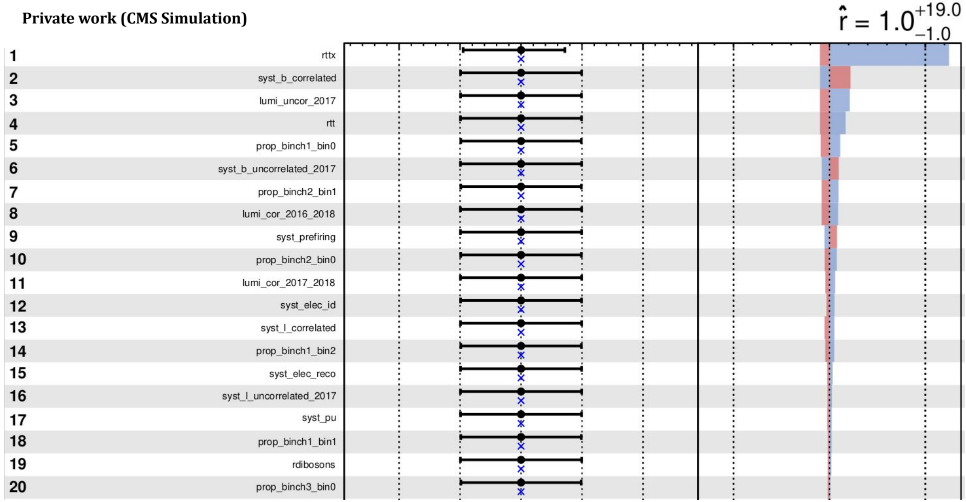


CR tt2l

Nuisance pools (2017)

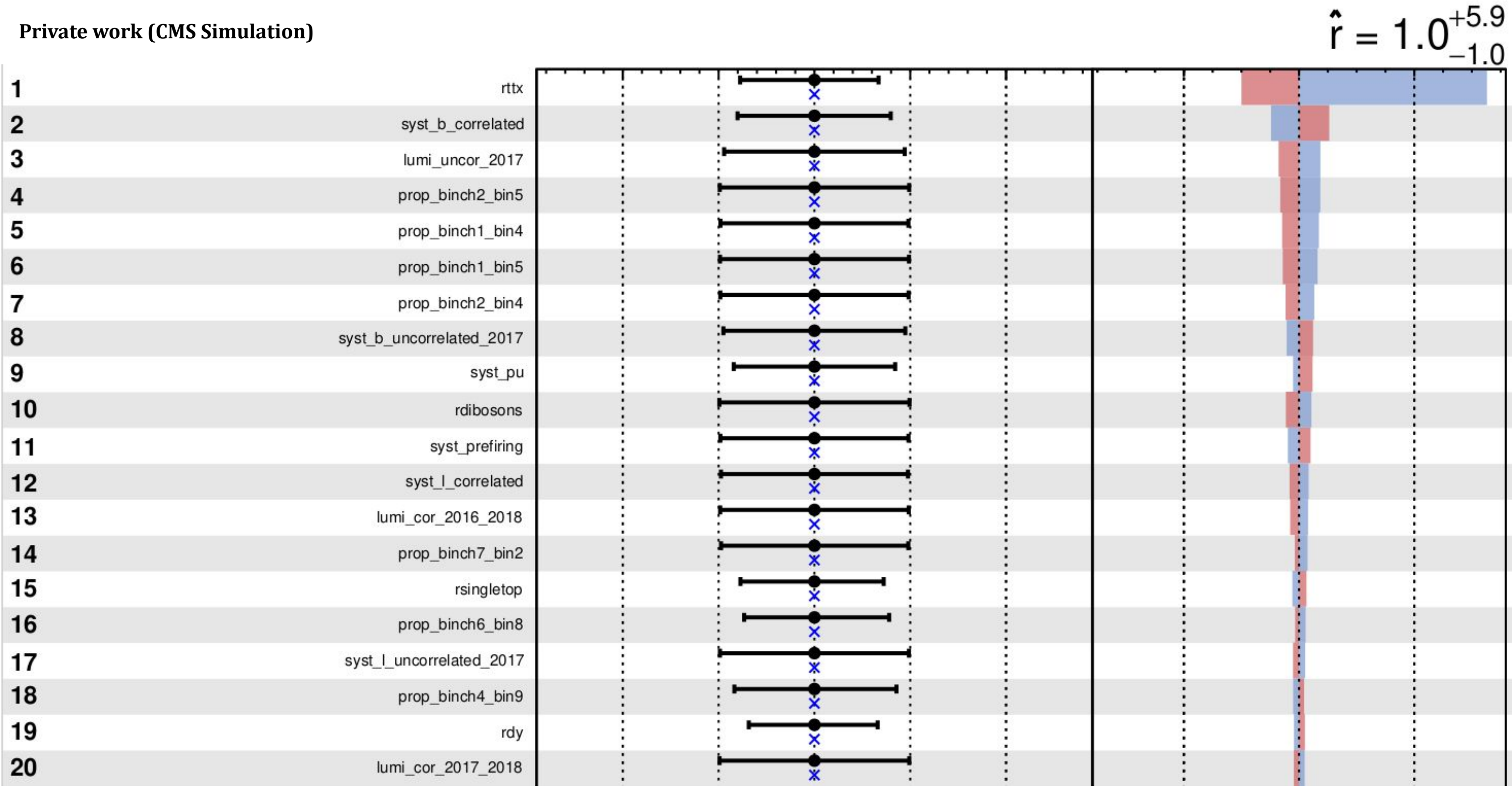


CR tt1l



CR ttX

Nuisance pools (2017)



All regions