Search for pair production of heavy particles decaying to a top quark and a gluon in the lepton+jets final state in proton-proton collisions at $\sqrt{s} = 13$ TeV at CMS

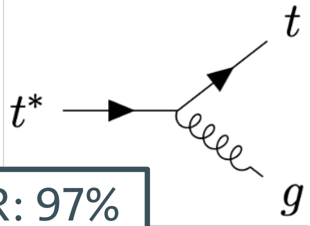
Finn Labe (Universität Hamburg)
on behalf of the CMS collaboration

arXiv:2410.20601 (submitted to EPJC)

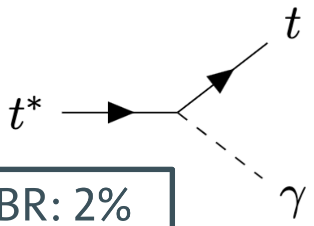
12.11.2024 | LHC TOP WG meeting

The particle in question: t*

- Top partner characterized by its decays:



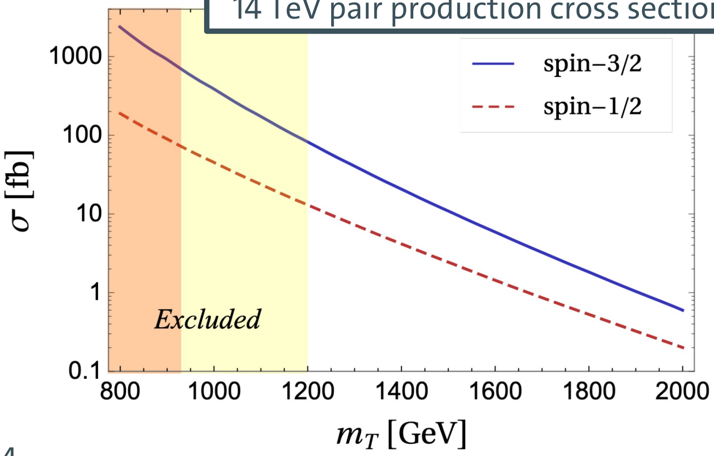
BR: 97%



BR: 2%

- Predicted in **composite top** models
- Same signature can appear for **VLQ** in little Higgs, extra dimensions…
  - Assuming **low mixing** with SM top
- This work: search for $t^*\bar{t}^* \rightarrow tg\bar{t}g$ (1ℓ)

$t\bar{t}$ + two jets

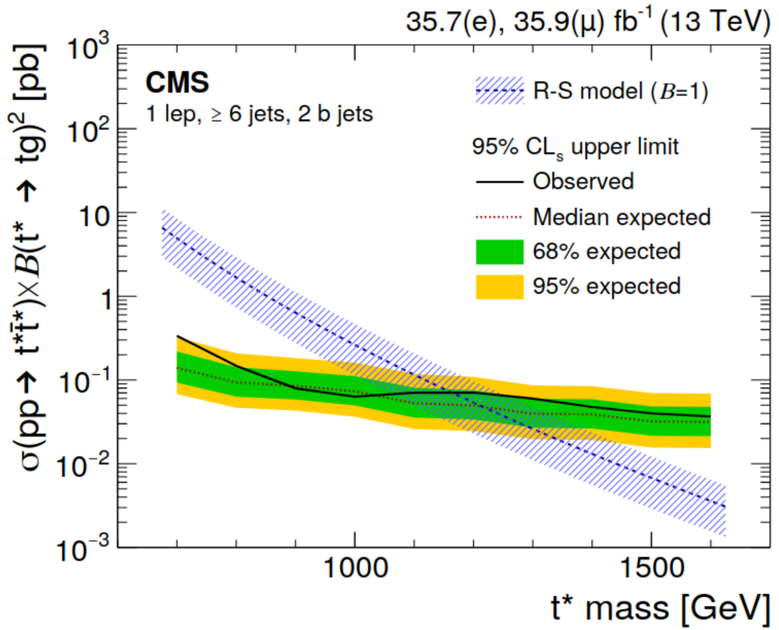


(H. Alhazmi et. al. *JHEP* 01 (2019) 139)

Past $t^*\bar{t}^* \rightarrow tg\bar{t}g$ (1ℓ) searches

*reinterpreting spin-3/2 results

- **Run 1 CMS analysis, 8 TeV**
  (CMS Collaboration, JHEP 06 (2014) 125.)
  - Excluded spin-1/2 t* up to 521 GeV*
  - Excluded spin-3/2 t* up to 803 GeV
- **Run 2 CMS analysis, 13 TeV** (2016 data)
  (CMS Collaboration. Phys.Lett.B 778 (2018) 349-370.)
  - t^* mass reconstruction approach
  - Assumed resolved jets from top quark
  - Excluded spin-3/2 t* up to 1200 GeV
- No ATLAS results on $t^*\bar{t}^*$ searches



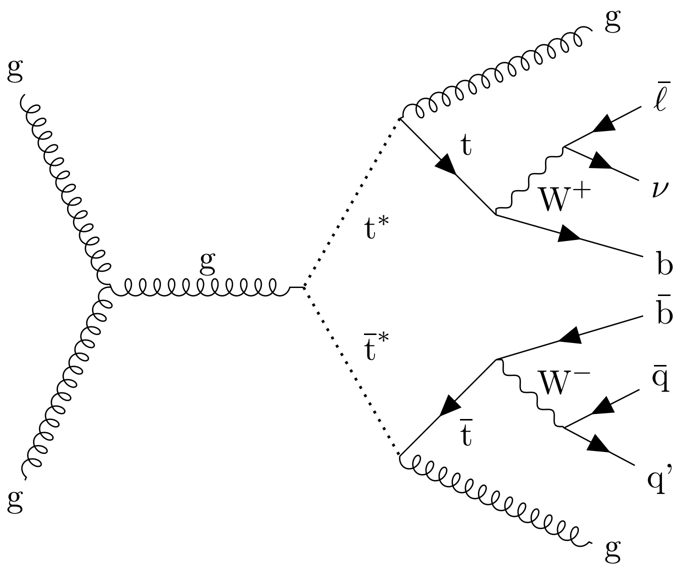
Current limits on a spin 3/2 t* by CMS
(CMS Collaboration. Phys.Lett.B 778 (2018) 349-370.)

Analyzed processes

estimated from data

- Signal processes: **spin-1/2 and spin-3/2 $t^*\bar{t}^*$**
  - Included in simulations with **EFT approach**:
    - Only added particle beyond SM: t^*
    - Couplings described by EFT operators
  - Main spin difference: total momentum
    - Difference gets smaller for higher m_{t^*}

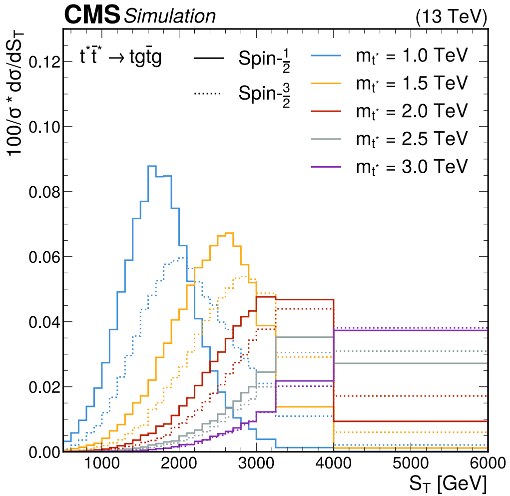
- SM backgrounds: $t\bar{t}$, W+jets, single top (ST), QCD, Diboson (VV) & Drell-Yan (DY)



Feynman diagram of $t^*\bar{t}^* \rightarrow tg\bar{t}g \rightarrow b\bar{b}q\bar{q}gg\ell\nu$

Search variable

- Previous analysis: mass reconstruction approach
  - Problem: many objects, difficult combinatorics
  - Alternative: simple variable, enrich signal with DNN



$$S_T = p_T^{\ell} + p_T^{\text{miss}} + \sum p_T^{\text{jets}}$$

Cut-based selection

SR definition with DNN

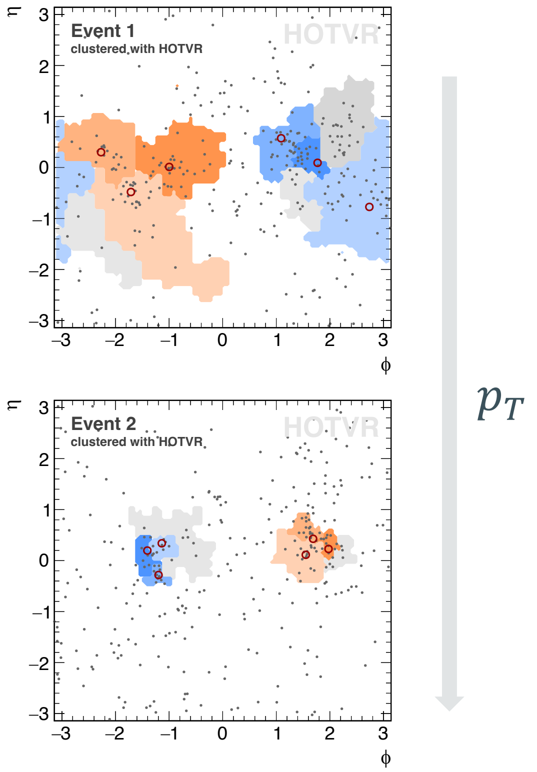
Background estimation

Cut-based event selection

- Standard **single lepton $t\bar{t}$** selection: Lepton, p_T^{miss}, jets (one b-tagged)

- Variable-radius jets used: HOTVR algorithm
  (T. Lapsien, R. Kogler, J. Haller, *Eur.Phys.J.C* 76 (2016) 11, 600.)
  - Reconstruct top quarks in a single jet for different top quark p_T
  - Provides **jet substructure** information
  - Can also reconstruct gluon jets, cover splitting

  → Good algorithm for this analysis!



(T. Lapsien, R. Kogler, J. Haller, *Eur.Phys.J.C* 76 (2016) 11, 600.)

Event classification DNN

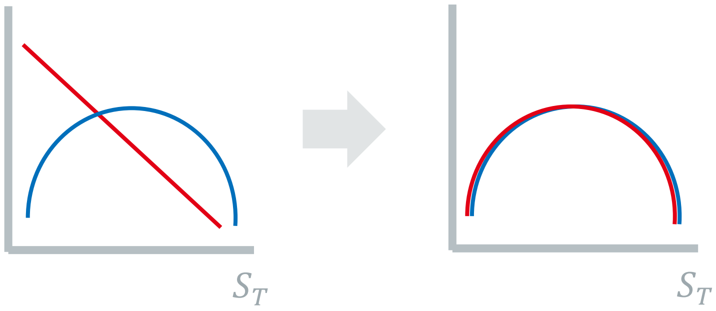
- Train DNN to discriminate signal from $t\bar{t}$
  - Outputs "signal-likeness score": s_{DNN}

33 input variables

- Leading lepton $p_T, \eta, \varphi, \mathrm{I}_{rel}$
- 3 HOTVR jets $p_T, \eta, \varphi, N_{\mathrm{subjets}}, \tau_{1,2,3}$
- 1 AK4 jet (highest b-tag)
  p_T, η, φ, DeepJet score
- $p_T^{\mathrm{miss}}, \vec{p}_T^{\mathrm{miss}}\ \varphi$
- $N_{\mathrm{AK4}}, N_{\mathrm{HOTVR}}$

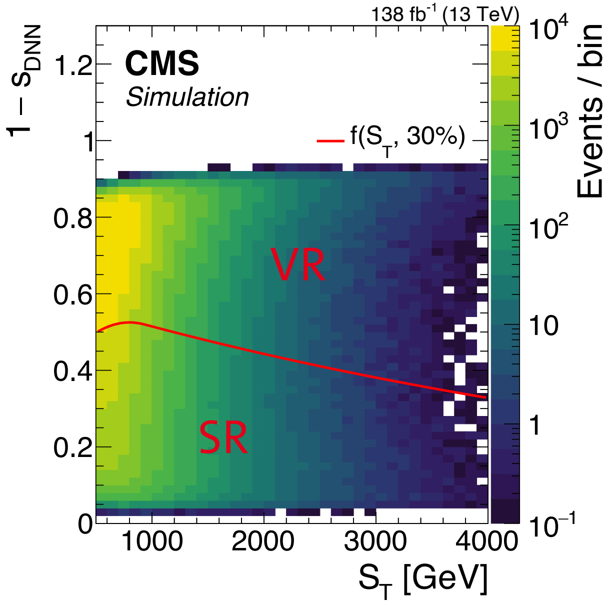
- Problem: many input variables strongly correlated with S_T

- **Decorrelation procedure** needed
  - Remove S_T difference between signal and $t\bar{t}$ with weights



Designing Decorrelated Taggers (DDT)

Removing remaining correlations by S_T-dependent DNN score threshold



$$s_{\mathrm{DDT}} = s_{\mathrm{DNN}} - f(S_T, \epsilon_{t\bar{t}})$$

Signal region (SR)

$s_{\mathrm{DDT}} > 0$

$t\bar{t}$ eff.:	30%
Signal eff.:	55 – 75%

Statistical analysis

Validation region (VR)

$s_{\mathrm{DDT}} < 0$

Low in signal, similar SM composition as SR

Validation

Analysis region overview

- Three main regions, each split into **muon** and **electron channel**
  - Control region: invert b-tagging requirement (no s_{DDT} condition)
  - **Only signal region used in statistical analysis**

Control region (CR)
b-jet veto

Low in signal, rich in non-t backgrounds

Background estimate

Signal region (SR)
$s_{\mathrm{DDT}} > 0$

$t\bar{t}$ eff.: 30%
Signal eff.: $55 - 75\%$

Statistical analysis

Validation region (VR)
$s_{\mathrm{DDT}} < 0$

Low in signal, similar SM composition as SR

Validation

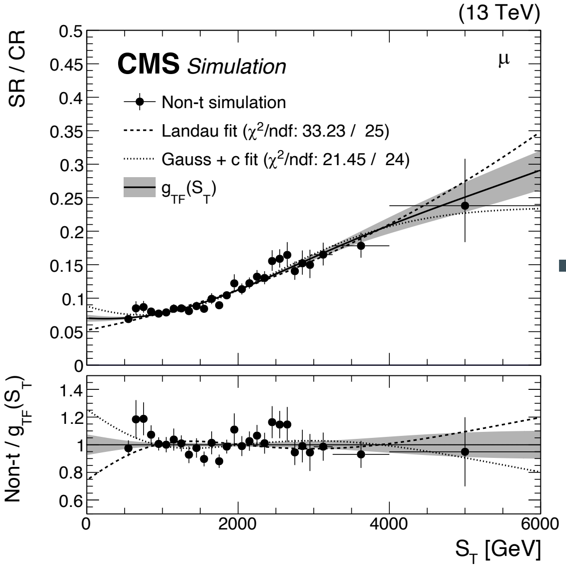
Data-driven background estimation

- Estimating non-top backgrounds from control region with transfer function

$$N_{SR}(\text{non}-\text{t bkg}) = g_{TF}(S_T)\big(N_{CR}(\text{data}) - N_{CR}(\text{top bkg})\big)$$

Control region
Enriched in non-t backgrounds with veto on b-jets

Subtract contribution from t-backgrounds

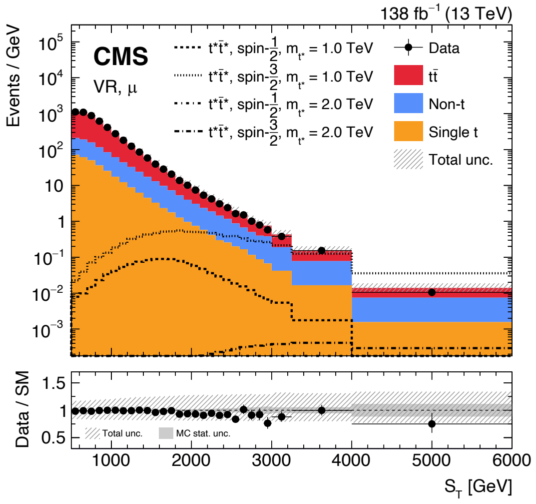


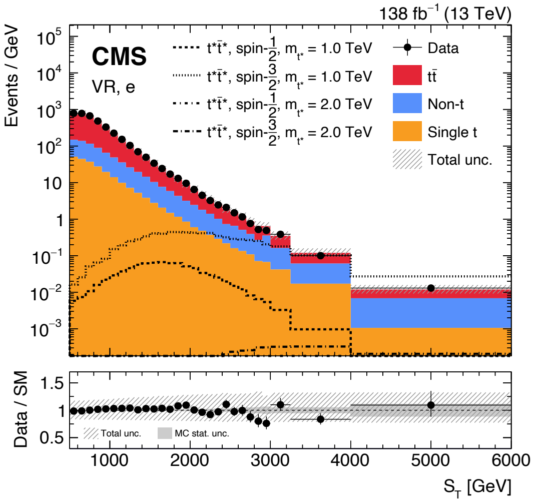
Signal region
non-t backgrounds

Validation region
non-t backgrounds

Background estimation: validation region

Pre-fit S_T distributions in low $S_{\mathrm{DDT}} < 0$ VR:





- Good data / SM – agreement in VR
- VR and SR are identical except for DNN cut
  - Nuisance accounts for any shape diff.

Signals scaled to theory prediction

Statistical analysis overview

- Binned **maximum likelihood fit** of both lepton channels simultaneously
  - Combining SR templates from all data-taking eras

- Including **nuisance parameters** for **systematic uncertainties** from
  - ... data / MC scale factors
  - ... background estimation
  - ... theory (μ_r, μ_f, top p_T reweighting, PDFs)
  - ...

Dominant experimental uncertainties:
b-tagging and DNN decorrelation

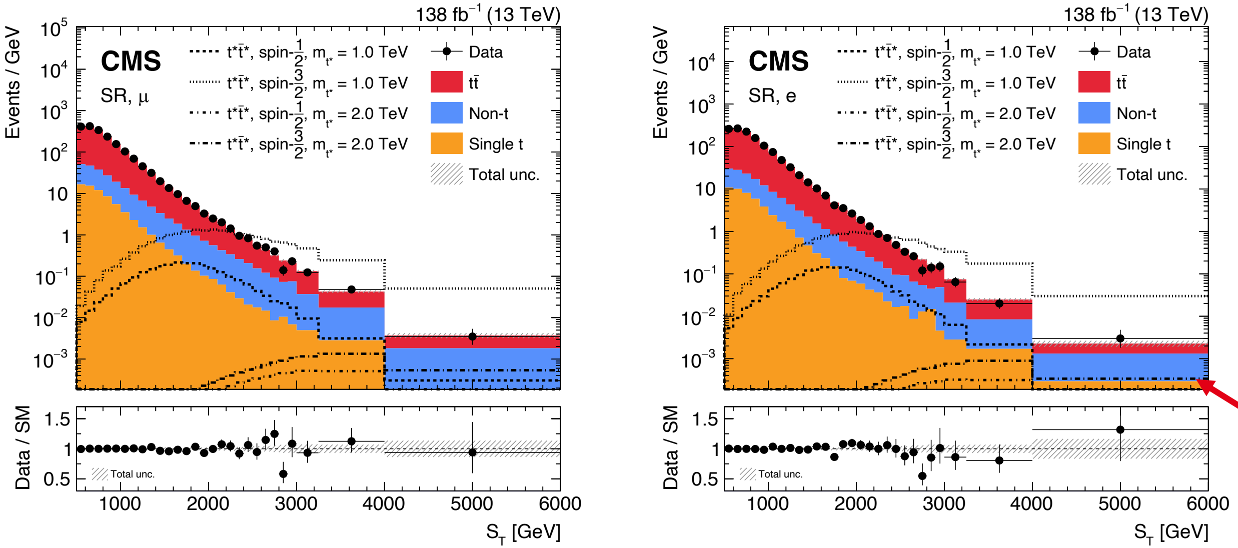
Dominant theoretical uncertainty:
Renormalization scale (μ_r) and factorization scale (μ_f) variations

Improved simulation accuracy could help here!

SR post-fit distributions

Signals scaled to theory prediction

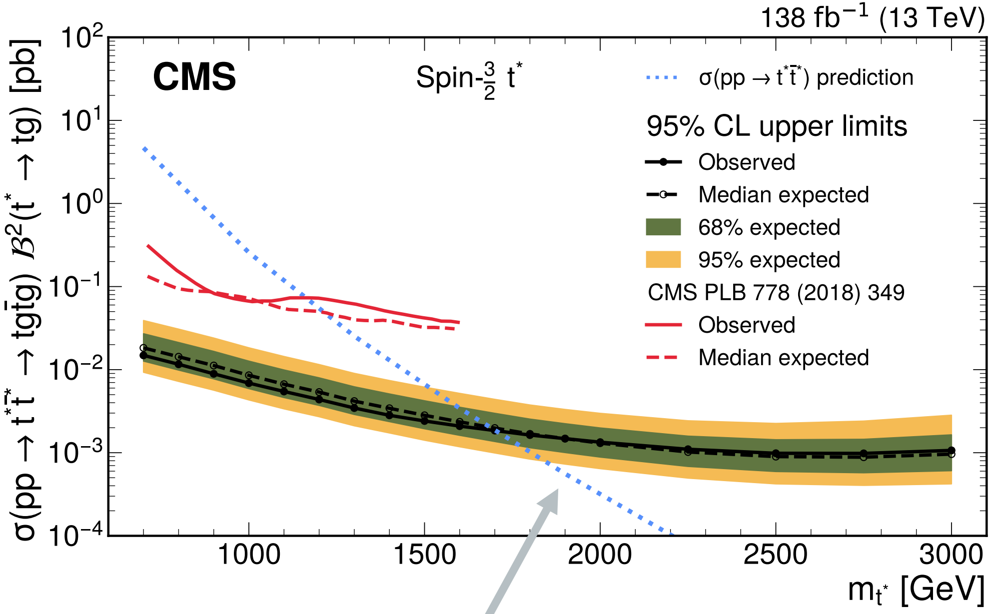
No deviation from SM predictions observed, background-only fit agrees well:



Large bins necessary due to low event count

Exclusion limits (spin 3/2)

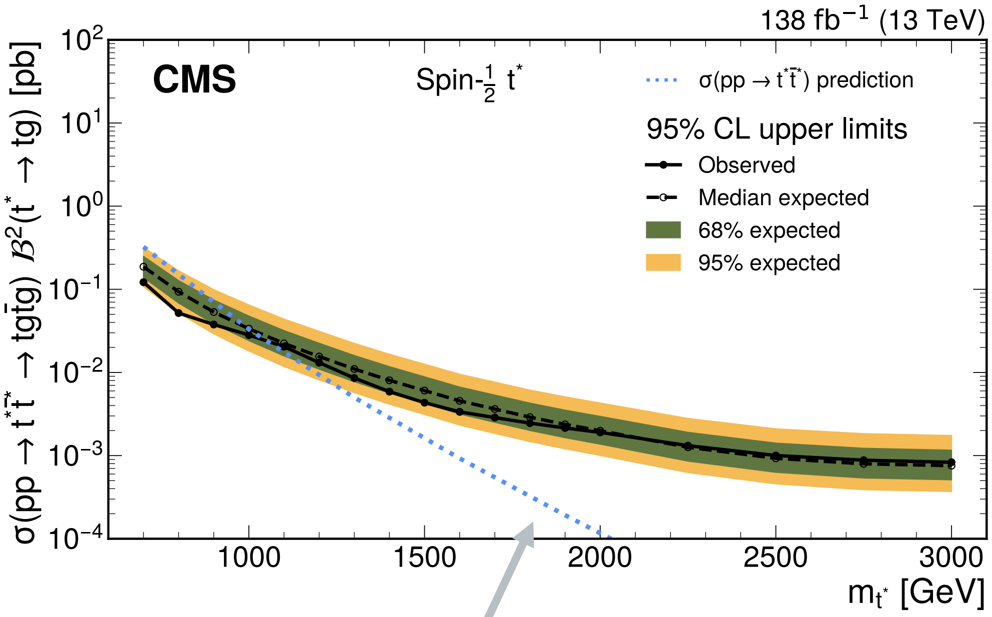
- Exclusion limits for spin 3/2 t*
- Significant improvement over 2016 CMS result
  - More data: $\sqrt{N} \approx 2$
  - Signal efficiency: x5
  - Background efficiency: x1.1
- Mass exclusion limit:
  - **1700 GeV observed**
  - 1690 GeV expected



Assuming 100% BR, theory predicts $(97\%)^2$

Exclusion limits (spin 1/2)

- Exclusion limits for spin 1/2 t*
- Worse than spin 3/2 limits, especially for low $m(t^*)$
- Mass exclusion limit:
  - **1050 GeV (observed)**
  - 990 GeV (expected)
- First ever 13 TeV exclusion limits for this spin scenario

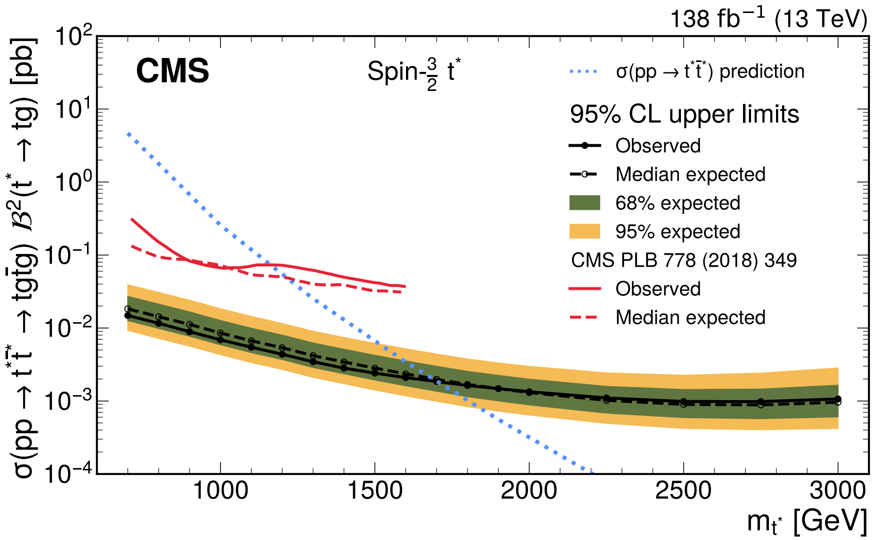


Assuming 100% BR, theory predicts $(97\%)^2$

Summary and outlook

- **Search for pair production of heavy quarks $t^*\bar{t}^* \rightarrow tg\bar{t}g$ (1ℓ)**
  - Substantially improved spin-3/2 limits, first ever 13 TeV spin-1/2 limits
- **Opportunities** for future analyses:
  - Larger dataset (Run 3, HL-LHC...)
  - Higher accuracy $t\bar{t}$ + 2 jets simulation
  - Analyzing the $t^*\bar{t}^* \rightarrow tg\bar{t}\gamma$ decay
  - Adding the t^*t process

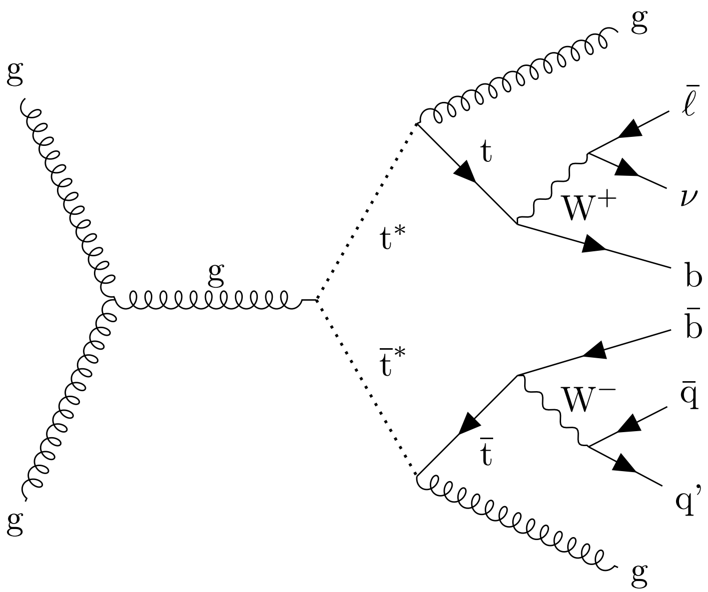
Cross section depends on coupling assumptions!



Backup

Cut-based event selection

- Single lepton trigger (μ or e)
- Exactly one lepton (μ or e)
- ≥ 4 small-radius (AK4) jets
- ≥ 1 variable-radius (HOTVR) jet
- $p_T^{\text{miss}} > 50$ GeV
- ≥ 1 medium DeepJet b-tag
- Custom lepton isolation
- $S_T > 500$ GeV



Feynman diagram of $t^*t^* \to tgtg \to bbqqgg\ell\nu$