Search for pseudoscalars and scalars decaying to top quark pairs with CMS Run 2

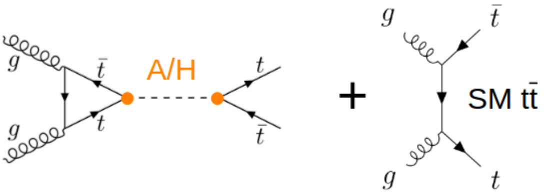
Samuel Baxter on behalf of the CMS Collaboration

Higgs Hunting 2024, Paris, 24.09.2024

Theoretical Motivation

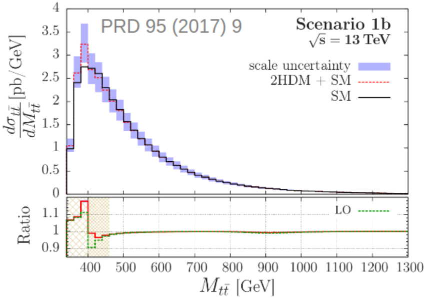
- Many extensions to the Standard Model rely on an extended Higgs sector
- Among the most popular SM extensions with an exteded Higgs sector are the 2HDM and MSSM
- We make the assumptions that the mass of the new (pseudo)scalar is more than twice the mass of the top quark, which it couples to exclusively

$$\mathcal{L}_A^{\text{int}} = i g_{\mathrm{At\bar{t}}} \frac{m_{\mathrm{t}}}{v} \bar{\mathrm{t}} \gamma_5 \mathrm{t} \mathrm{A} \qquad \mathcal{L}_H^{\text{int}} = -g_{\mathrm{Ht\bar{t}}} \frac{m_{\mathrm{t}}}{v} \bar{\mathrm{t}} \mathrm{t} \mathrm{H}$$



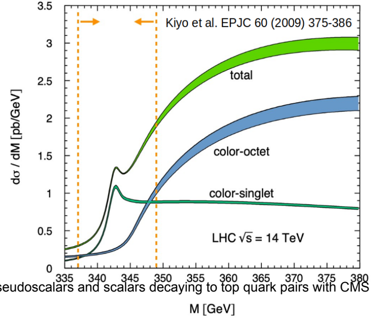
Signal Model

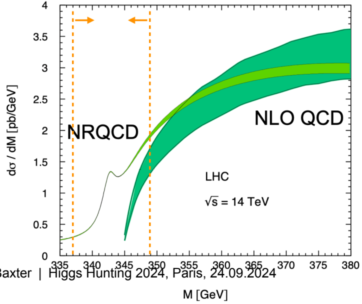
- In our search, we use a model-independent approach, where we consider the mass, CP parity, width and the coupling modifiers $g_{\mathrm{A}}/g_{\mathrm{H}}$ as free parameters for the new scalar/pseudoscalar
- In addition to the $\mathrm{t\bar{t}}$ resonance, the interference with Standard Model $\mathrm{t\bar{t}}$ production is also considered, resulting in a peak-dip structure in the $m_{\mathrm{t\bar{t}}}$ distribution
- The signal samples are generated at LO with NNLO QCD k-factors applied



Model of the $t\bar{t}$ Bound State (η_t)

- At the $m_{t\bar{t}}$ threshold, non-perturbative QCD effects start to play a role
- These effects inculde a bound state of $t\bar{t}$, appearing as a pseudoscalar resonance
- We model this effect with the following assumptions and parameters suggested by Fuks et al. (PRD 104 (2021) 3, 034023):
  - Color-singlet pseudoscalar particle η_t coupling to gluons and top quarks
  - $m_{\eta_t} = 343$ GeV, $\sigma_{\eta_t} = 6.43$ pb
  - Generated in MadGraph, filtering out events outside $m_{\eta_t} = 343 \pm 6$ GeV





Selections

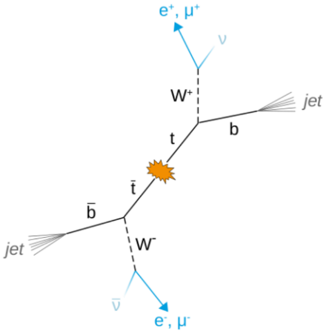
Dileptonic Channel

Two oppositely charged leptons (ee/eμ/$\mu\mu$)

At least two jets

At least one jet b tagged

At least one solution to the $t\bar{t}$ reconstruction is required

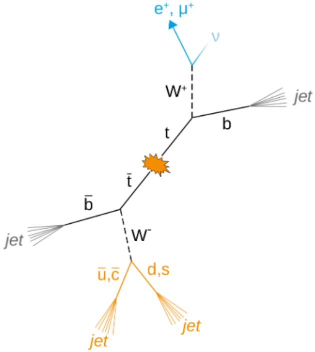


Semileptonic Channel

Only one lepton (e/μ)

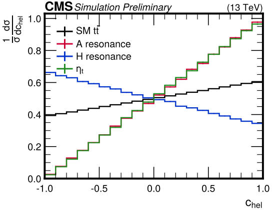
Three or more jets

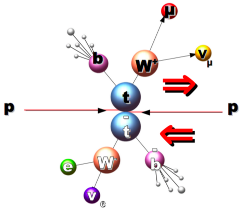
At least two jets b tagged

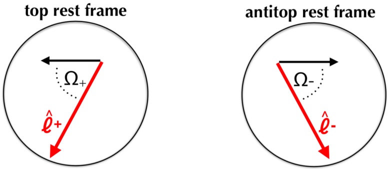


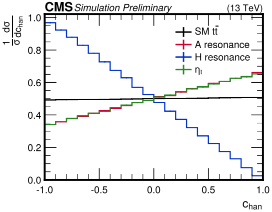
Dileptonic Channel

- The search variables are $m_{t\bar{t}}$, c_{hel} and c_{han}
- The c_{hel} variable is defined by taking the scalar product of the lepton unit vectors, boosted to the zero mass frame of the $t\bar{t}$ system and further boosted to their respective parent top quark
- The c_{han} variable has a similar definition, just has a flipped sign for the coordinate in the direction of the parent top quark for one of the leptons









Semileptonic Channel Variables

- The search variables in the semileptonic channel are $m_{t\bar{t}}$ and $|\cos(\theta^*)|$
- θ^* is the angle between the leptonically decaying top quark in the zero mass frame of the $t\bar{t}$ system and the diraction of the $t\bar{t}$ system in the laboratory frame
- For SM $t\bar{t}$ production, $|\cos(\theta^*)|$ peaks at $|\cos(\theta^*)| = 1$, whereas both the scalar and pseudoscalar signals are flat in comparison

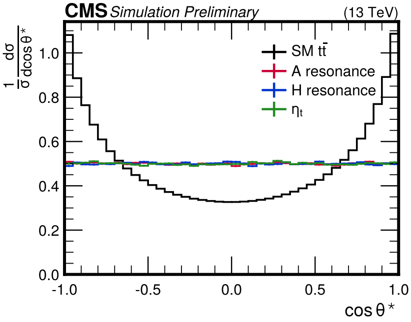
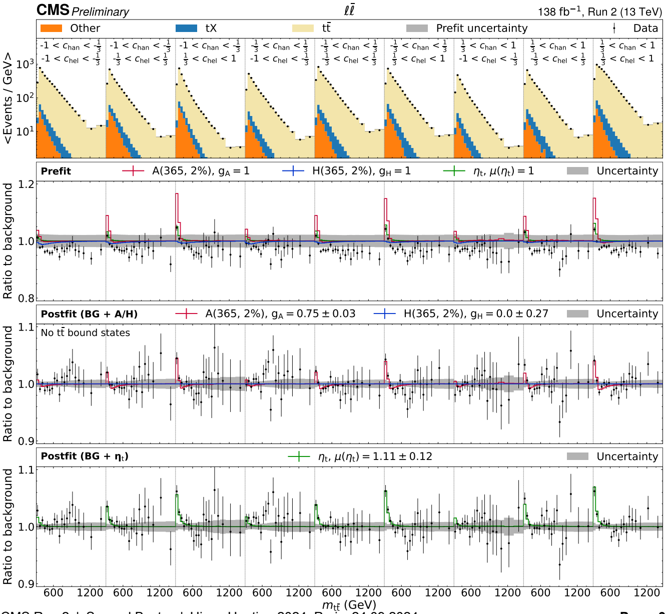


Figure: Gen-level distribution of the spin-correlation variable $\cos(\theta^*)$.

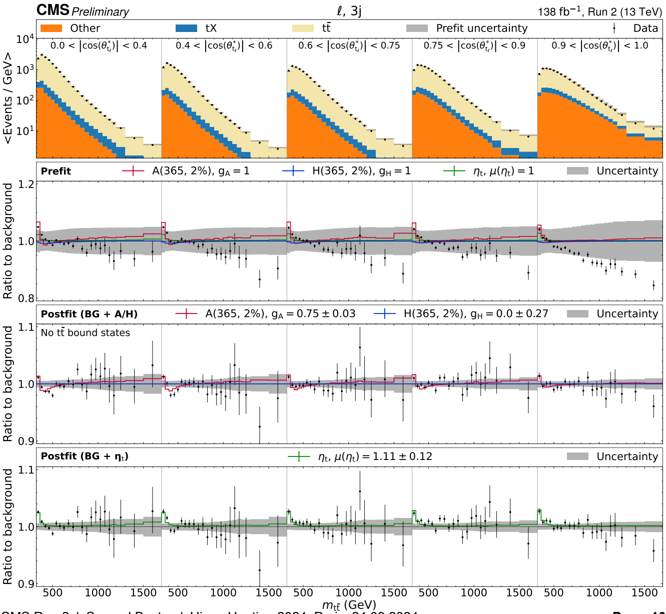
Backgrounds

- **SM $t\bar{t}$** is an irreducable background - the best possible precision and accuracy is needed
  - Generated at NLO
  - Shape reweighted to NNLO QCD + NLO EW, normalized to NNLO + NNLL
- **Single top and tW**, generated at NLO, present in both the single leptonic and dileptonic searches
- **Z+jets**, present in the dileptonic search, mainly the ee and $\mu\mu$ channels
  - MiNNLO samples for better data/MC agreement at low m_{ll}
  - Normalisation from Z window using an R_{in}/R_{out} method
- **QCD and EW**, data-driven estimate - semileptonic only
  - Sideband regions, where no jets pass the b-tagging
  - Transfer factors on W+jets and QCD MC simulations, shapes are corrected as well
- tt+X and Diboson MC simulations are included in the dileptonic search

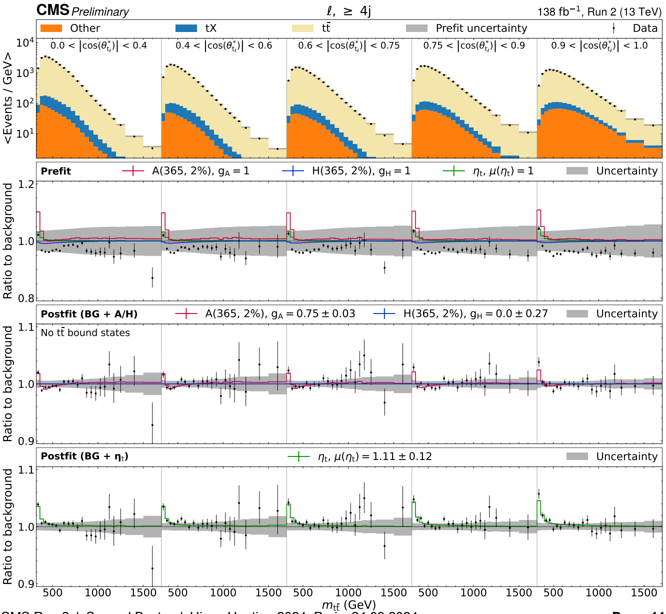
Prefit/Postfit Dilepton



Prefit/Postfit Lepton+3Jets

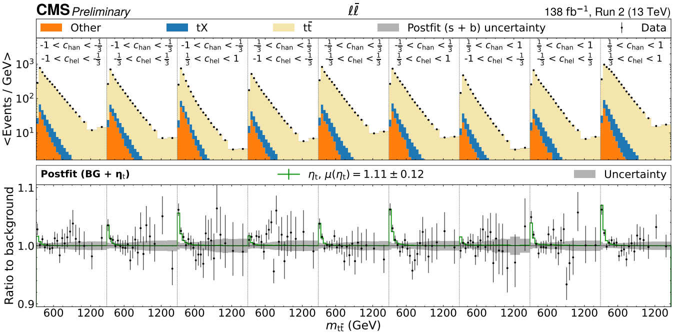


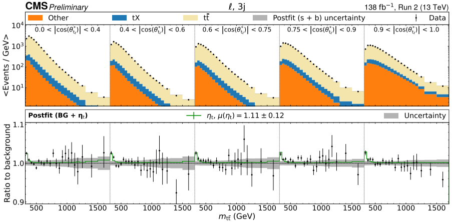
Prefit/Postfit Lepton+$\geq$4Jets

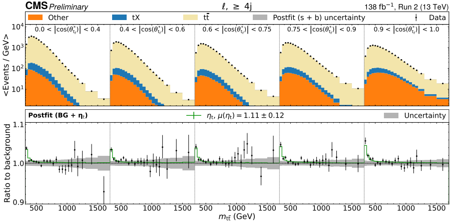


Best fit for η_t

- As seen in the postfit plots with the SM+η_t interpretation, data fits η_t well, and a cross section of **7.1 pb±11 %** is determined

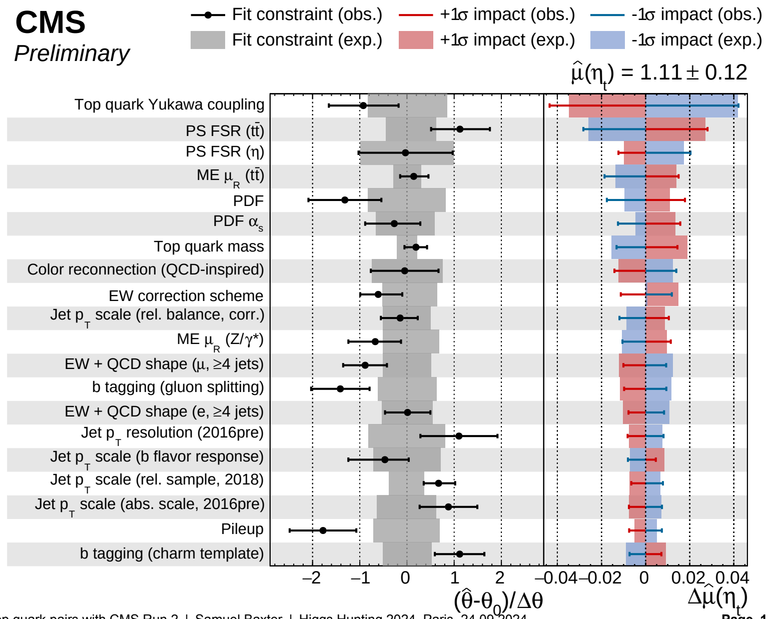






Uncertainties

> Modelling uncertainties have the dominating impact on fitting the η_t model



Combined Limit

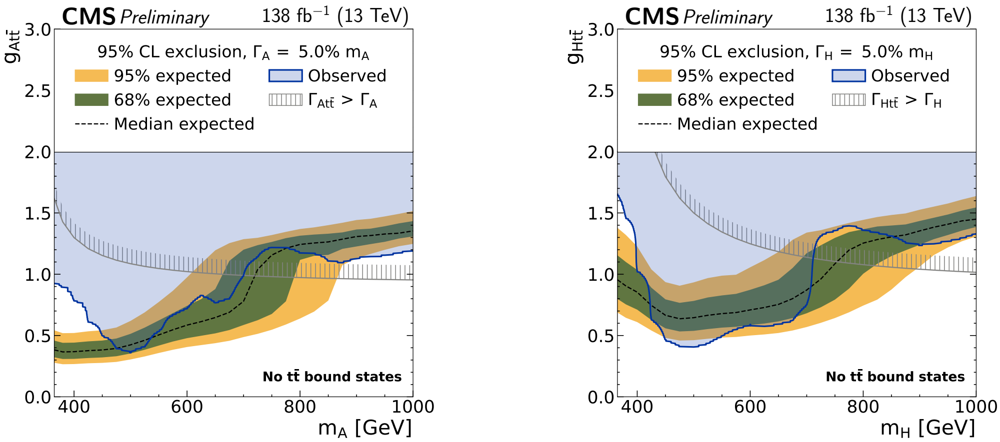


Figure: Limits for pseudoscalar(left) and for scalar(right), perturbative QCD SM background

Combined Limit with η_t

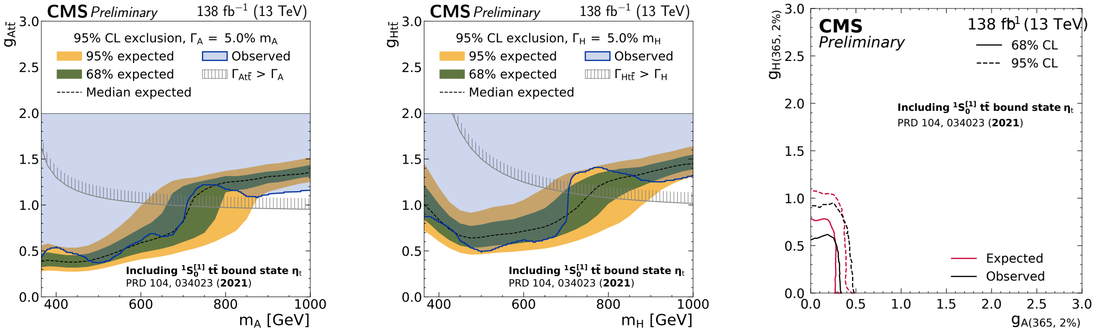


Figure: Limits for pseudoscalar(left), for scalar(centre) and 2-dimensional for both scalar and pseudoscalar (right), SM background with a model of the $t\bar{t}$ bound state

Summary

- **A search for a scalar or pseudoscalar has been peformed** in the dileptonic and semileptonic final states of $\mathrm{t\bar{t}}$ using the full Run2 dataset of CMS
- The analysis targets the $m_{\mathrm{t\bar{t}}}$ distribution along with angular and spin observables
- **A local excess, at low $m_{\mathrm{t\bar{t}}}$, has been observed in data with >5σ significance**, which fits better to **pseudoscalar** than scalar hyptheses
- Excess also fits best to a model of **the $\mathrm{t\bar{t}}$ bound state η_{t}**, a cross section has been determined
- **Stringent limits have been set** on the scalar and pseudoscalar signal models, with a floating normalisation of η_{t} included

Reference: CMS-PAS-HIG-22-013

Thank you!

Contact

Deutsches Elektronen-
Synchrotron DESY

www.desy.de

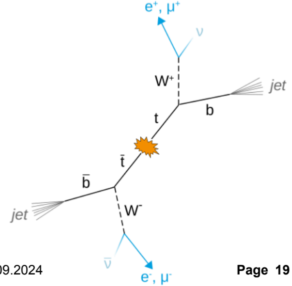
Samuel Baxter
0009–0008–4191–6716
CMS Exotics
samuel.baxter@desy.de
+49–40–8998–2008

Backup

Dileptonic Channel Selection

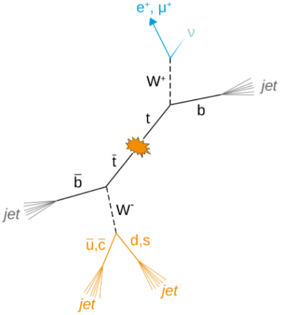
- **Two oppositely charged leptons** with $p_T > 25(20)$ GeV for the leading (subleading) lepton, $|\eta| < 2.4$, Split into three channels: ee eμ and $\mu\mu$
  - Muon: tight cut-based ID & ISO, Electrons: tight WP of MVA+ISO ID
  - Remove $m_{ll} < 20$ GeV, Z-window ($M_Z \pm 15$ GeV), along with a $p_T^{miss} > 40$ GeV requirement, in ee/$\mu\mu$
  - Pass any single lepton or dilepton trigger and veto additional leptons with $p_T > 20$ GeV
- **At least two jets** with $p_T > 30$ GeV, $|\eta| < 2.4$, $\Delta R > 0.4$ to leptons, tight ID
- **At least one jet b tagged** with a likelihood-based method, medium WP

- A reconstruction of the $t\bar{t}$ system is done analytically, **at least one solution to the reconstruction is required**
  - Assigning p_T^{miss} to the two neutrinos, assuming on-shell tops and Ws
  - Selecting the best fitting jets to the m_{lb} systems
  - Smeared masses, use weighted average of 100 variations

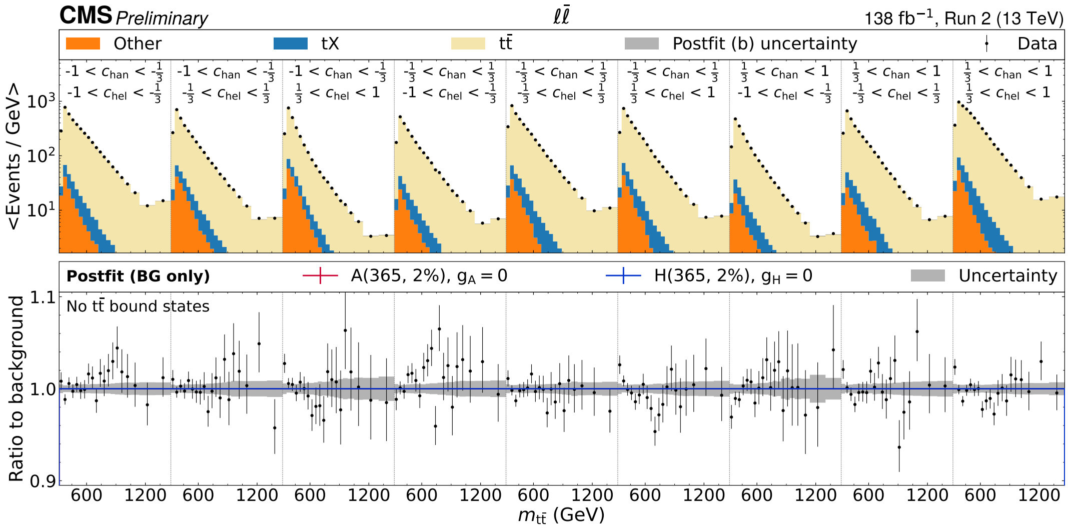


Semileptonic Channel Selection

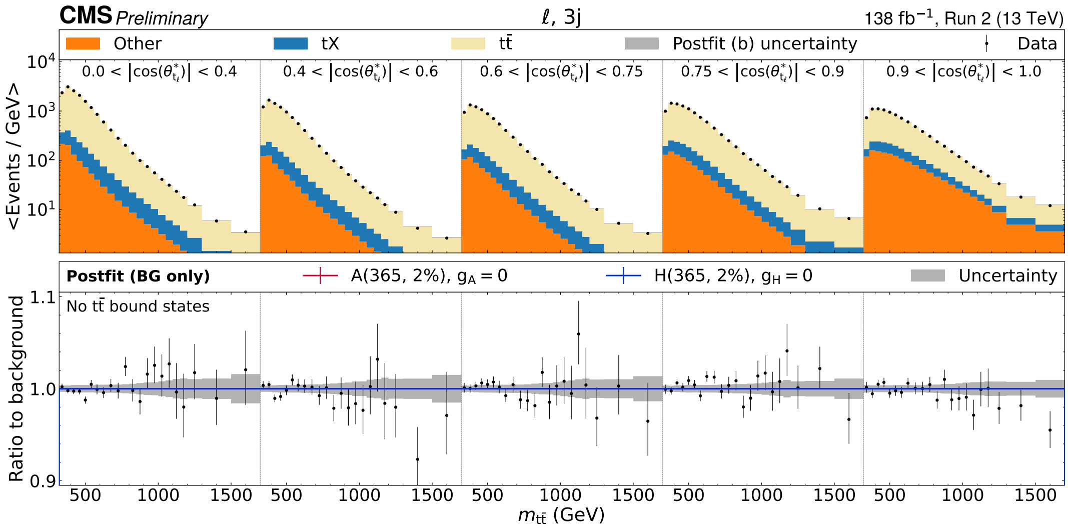
- **Require only one lepton (e/μ)** with $p_T > 30$ GeV, $|\eta| < 2.4$ with tight cut-based ID
  - Pass a single lepton trigger, veto additional leptons with $p_T > 20$ GeV and loose ID
  - For muon: tight PF ISO
- **Three or more jets**: $p_T > 30$ GeV, $|\eta| < 2.4$, $\Delta R > 0.4$ to leptons, tight ID
- **At least two jets b tagged** with a likelihood based method, medium WP
- Split in four channels: e vs μ and 3 jets vs ≥ 4 jets
- The $t\bar{t}$ system is reconstructed using an algorithm that applies:
  - B-jet assignment by maximum likelihood
  - Energy correction factor for 3 jet events to compensate for a missing jet in the event



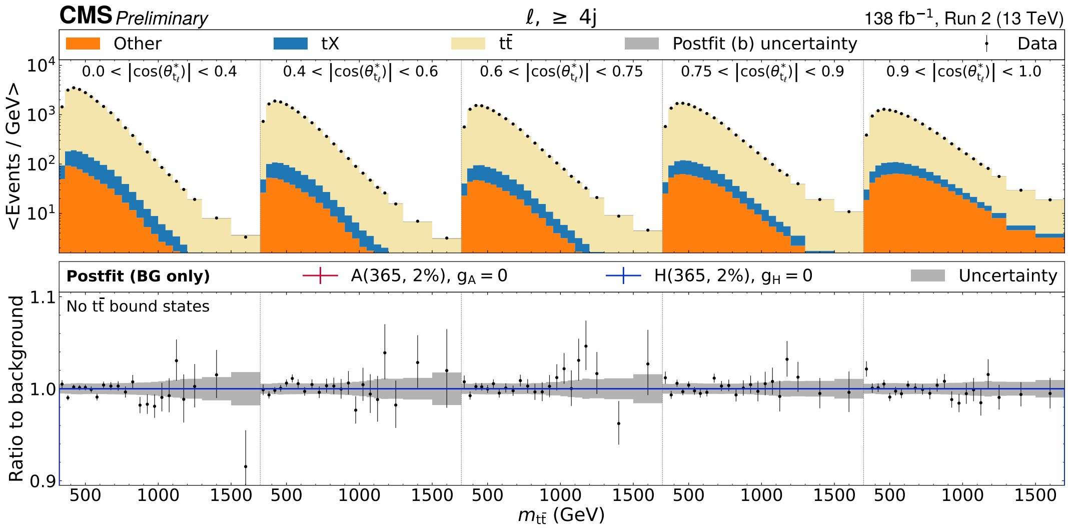
Postfit Pertubative SM Background Dilepton



Postfit Pertubative SM Background Lepton+3Jets

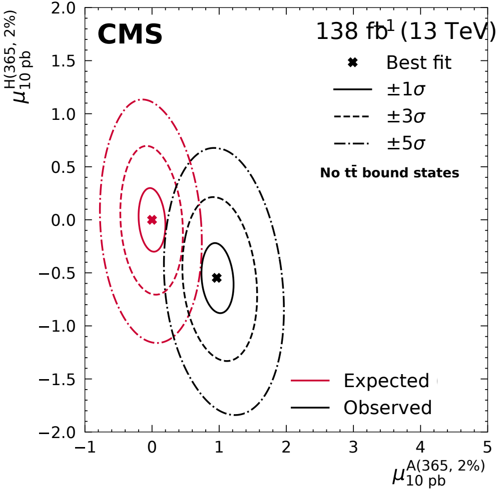


Postfit Pertubative SM Background Lepton+$\geq$4Jets



Testing the Result: Scalar vs Pseudoscalar

- To test whether the result fits better with a scalar or pseudoscalar, a simultaneus 2D fit is performed
- The fitted signals are resonant components of H and A with a mass of 365 GeV and a width of 2 %
- both cross sections are normalized to the same value (10 pb)
- As seen in the figure, data favours the pseudoscalar hypothesis



Hypothesis test

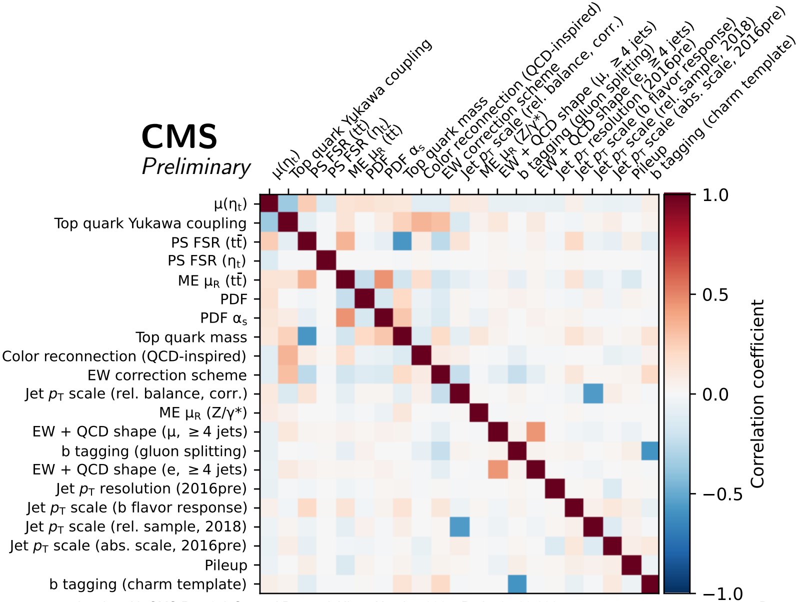
	Best-fit point	Difference in $-2\ln L$
η_t interpretation	$\mu(\eta_t) = 1.11$	-86.2
Single A interpretation	$m_A = 365\,\text{GeV}, \Gamma_A/m_A = 2\%, g_{At\bar{t}} = 0.78$	-72.6
Single H interpretation	$m_H = 365\,\text{GeV}, \Gamma_H/m_H = 2\%, g_{Ht\bar{t}} = 1.45$	-10.4

- When comparing the best fitting generated scalar and pseudoscalar signal ($m_A = 365$ GeV, $\Gamma_A = 2\%$) to SM with our η_t toy model, data prefers η_t
- One needs to keep in mind that signal samples were only generated down to $m_{A/H} = 365$ GeV so far, and that interference is included

Table B.1: Results on the η_t cross section, using only the $\ell\bar{\ell}$ channels, for the b_bbar_4l background prediction and for the default setup. The quoted uncertainty for b_bbar_4l assumes the same uncertainty as for the nominal result.

Prediction for SM $t\bar{t}$ and tW	Extracted η_t cross section	Uncertainty
b_bbar_4l (POWHEG vRES)	5.9 pb	18%
Default (POWHEG v2)	7.5 pb	13%

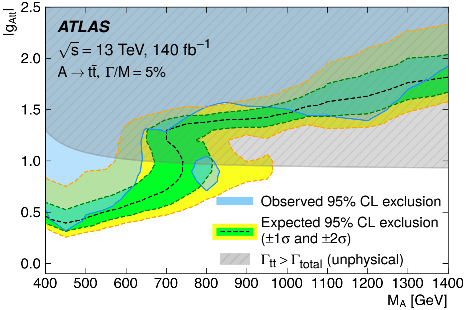
Correlation Matrix



Related Searches

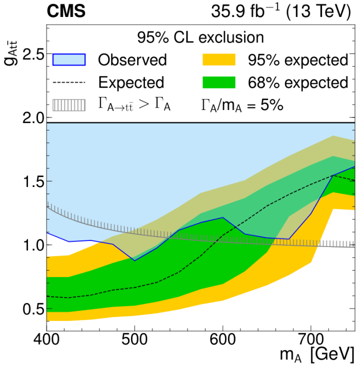
Recent full Run 2 result by ATLAS

JHEP 08 (2024) 013



Previous CMS A/H search:
CMS-HIG-17-027 (35.9 fb^{-1})

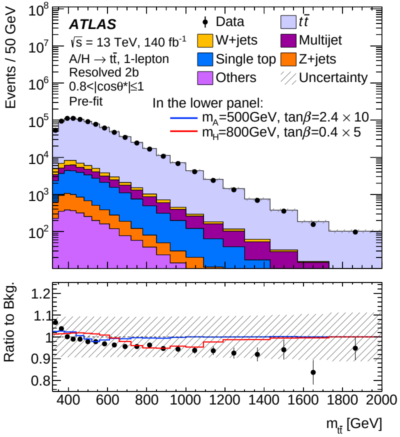
JHEP 04 (2020) 171



ATLAS Prefits

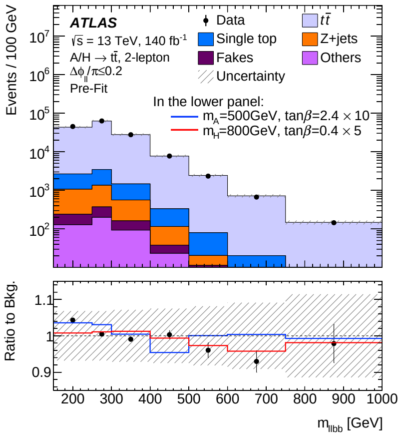
ATLAS prefit example lj channel

JHEP 08 (2024) 013



ATLAS prefit example ll channel

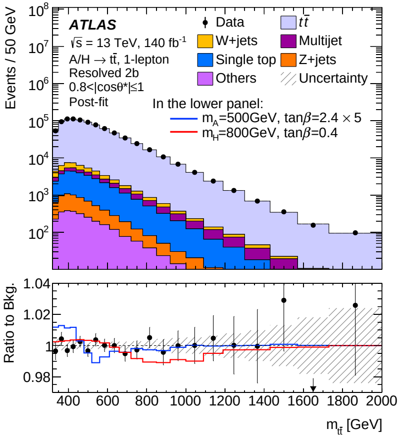
JHEP 08 (2024) 013



ATLAS Postfits

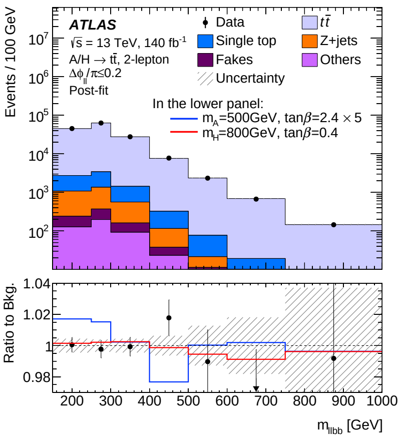
ATLAS postfit example lj channel

JHEP 08 (2024) 013



ATLAS postfit example ll channel

JHEP 08 (2024) 013



Combined Limit

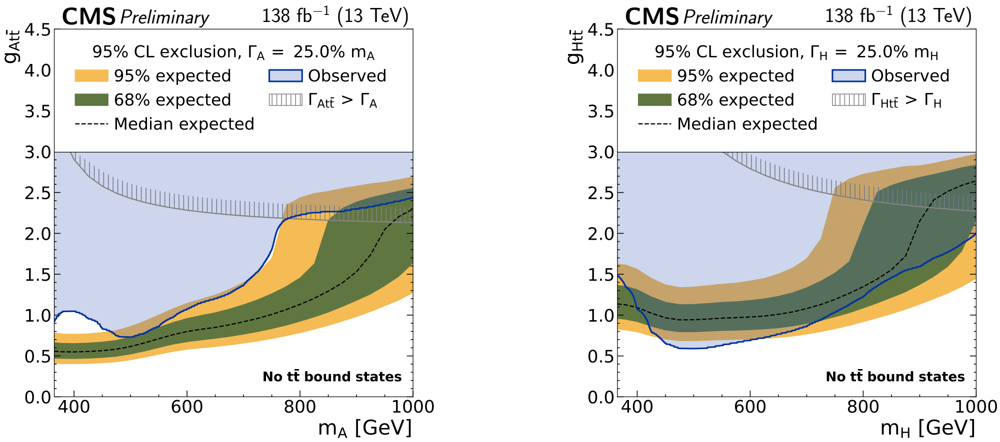
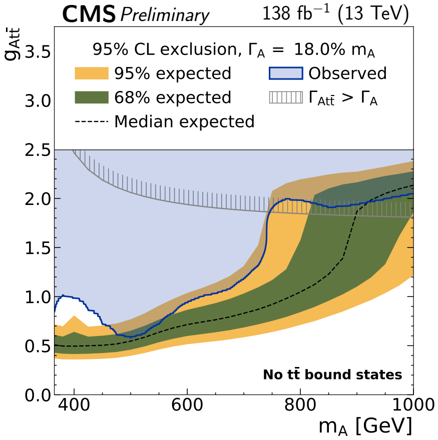


Figure: Limits for pseudoscalar(left) and for scalar(right), perturbative QCD SM background

Combined Limit



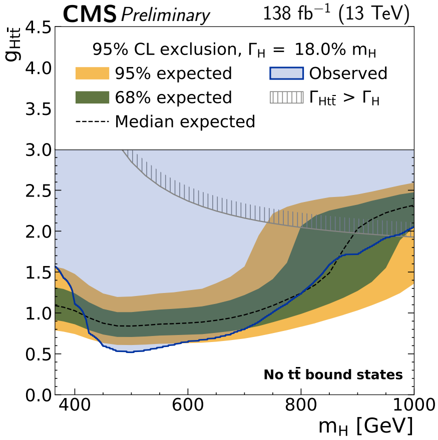
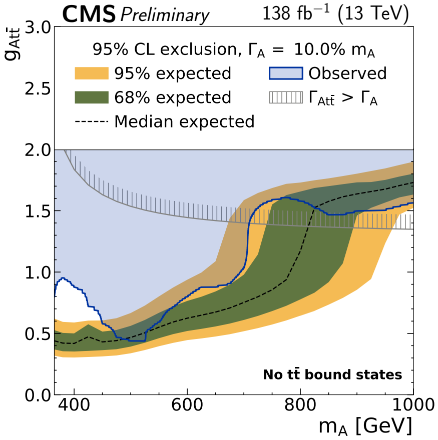


Figure: Limits for pseudoscalar(left) and for scalar(right), perturbative QCD SM background

Combined Limit



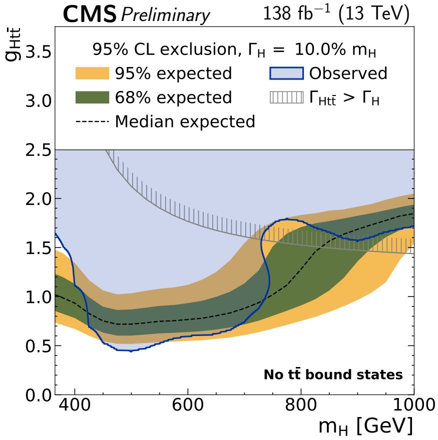


Figure: Limits for pseudoscalar(left) and for scalar(right), perturbative QCD SM background

Combined Limit

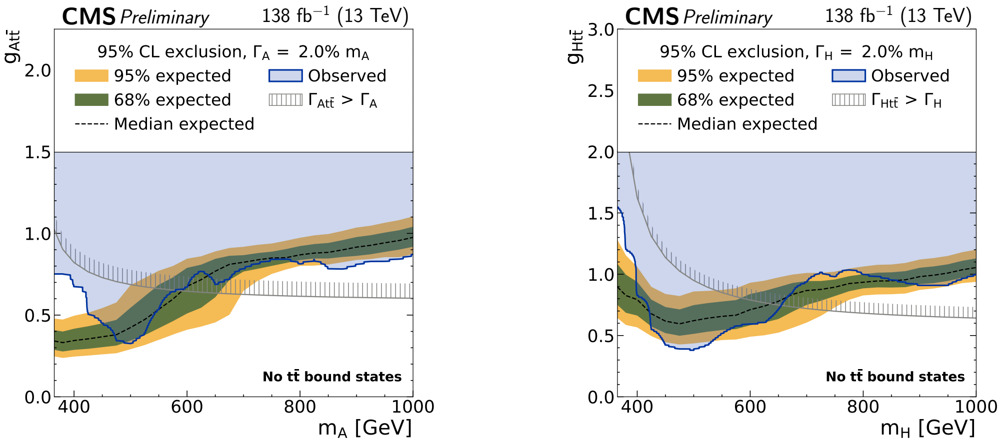


Figure: Limits for pseudoscalar(left) and for scalar(right), perturbative QCD SM background

Combined Limit

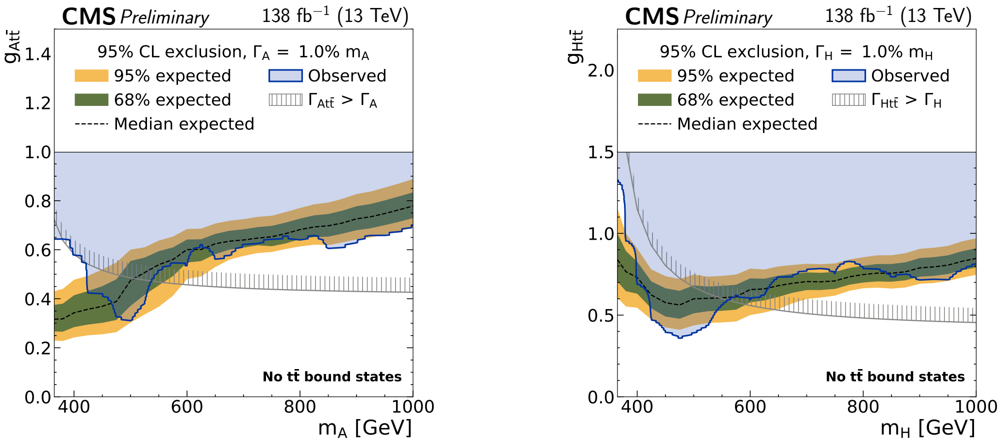
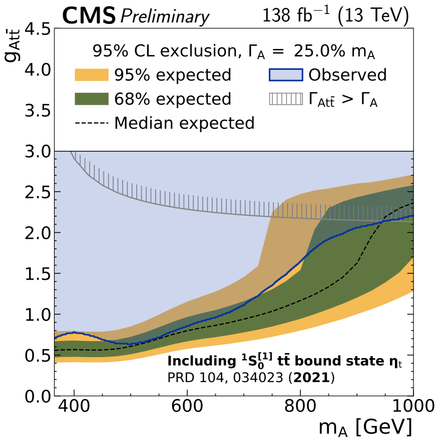


Figure: Limits for pseudoscalar(left) and for scalar(right), perturbative QCD SM background

Combined Limit with η_t



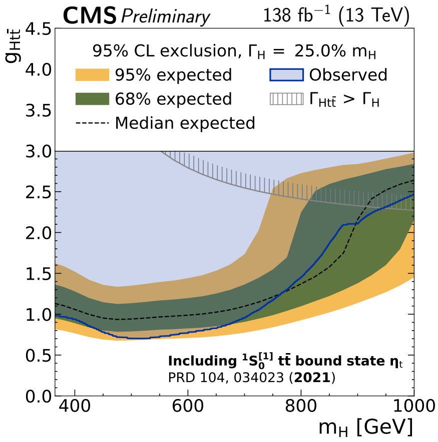
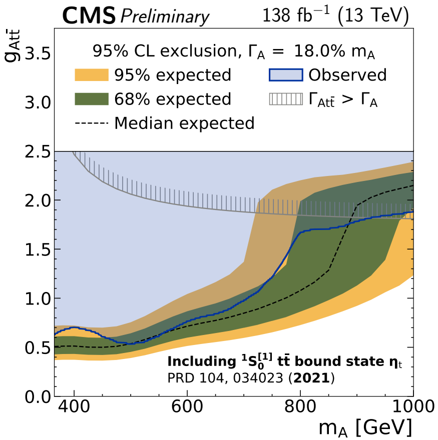


Figure: Limits for pseudoscalar(left) and for scalar(right), SM background with non-perturbative toy-model

Combined Limit with η_t



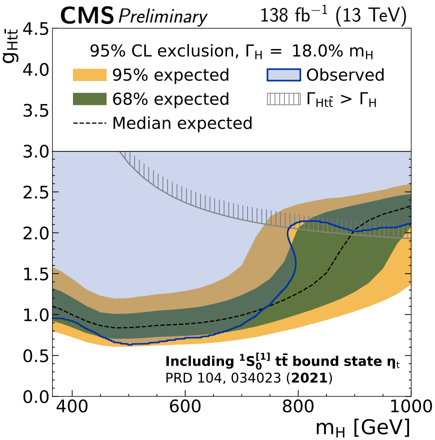
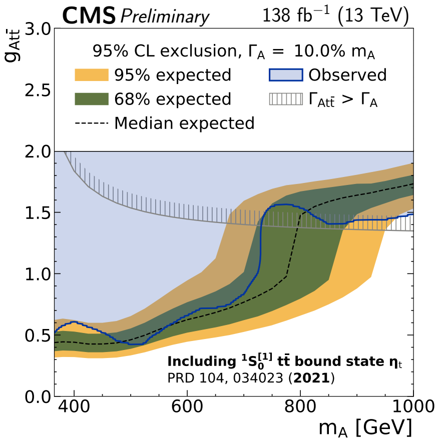


Figure: Limits for pseudoscalar(left) and for scalar(right), SM background with non-perturbative toy-model

Combined Limit with η_t



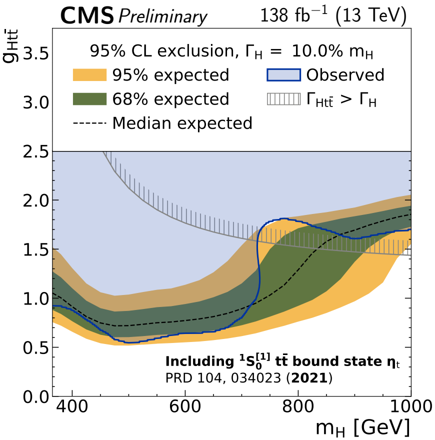


Figure: Limits for pseudoscalar(left) and for scalar(right), SM background with non-perturbative toy-model

Combined Limit with η_t

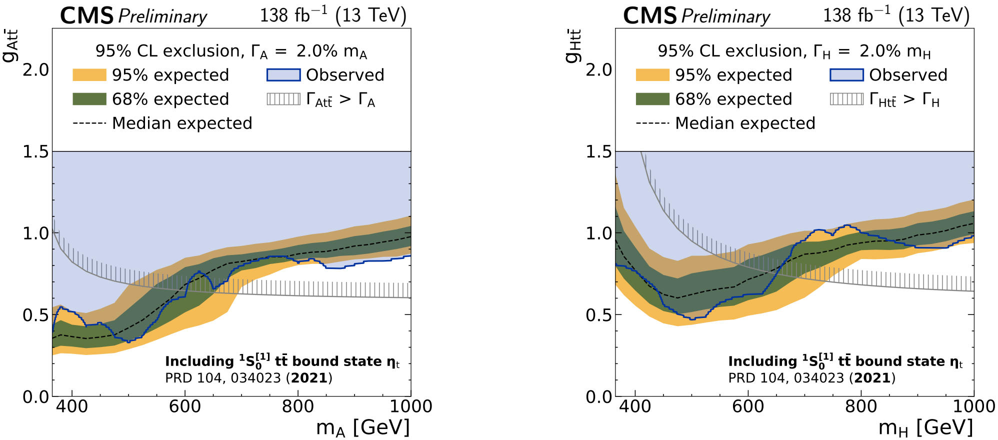
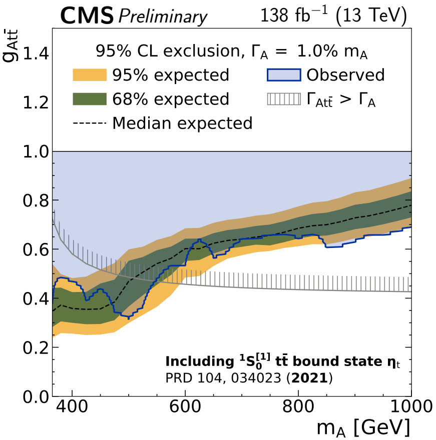


Figure: Limits for pseudoscalar(left) and for scalar(right), SM background with non-perturbative toy-model

Combined Limit with η_t



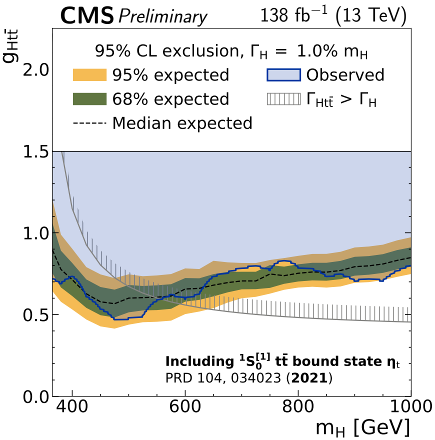


Figure: Limits for pseudoscalar(left) and for scalar(right), SM background with non-perturbative toy-model

Combined contours with η_t

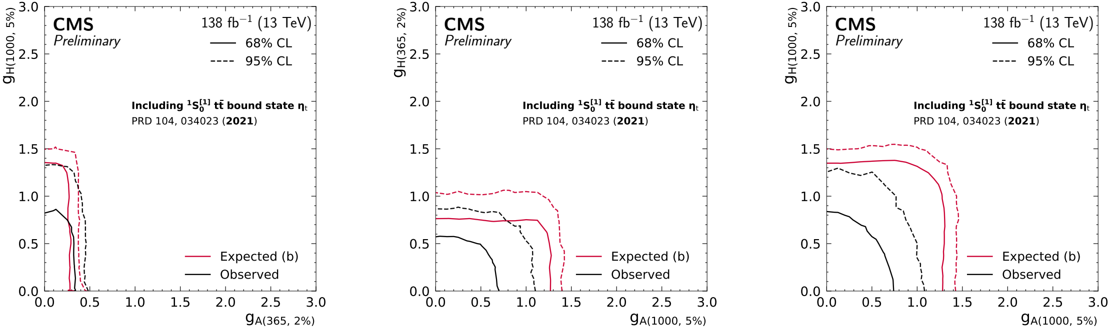


Figure: Contours for pseudoscalar and for scalar simultaneously, SM background with non-perturbative toy-model