LHCTopWG Open Meeting

Search for heavy scalar or pseudoscalar states in $t\bar{t}$ events at CMS

Laurids Jeppe for the CMS collaboration

13.11.2024 | CMS-PAS-HIG-22-013



Motivation: BSM!

- We know there is BSM but where is it?
- Many models predict extended Higgs sectors e.g. 2HDM, MSSM…
 → new scalar or pseudoscalar states
- If couplings are Yukawa-like: strongest coupling to top quark
- If mass > 2m_t: decay to tt
 is dominant in large areas of parameter space

 \rightarrow search for heavy (pseudo)scalars in tt final states





Motivation: tt bound states?

SM predicts tt (quasi-)bound states below the tt threshold



- Not observed yet (but there are hints: differential tt, entanglement...)
- Dominant component: pseudoscalar can we search for it?

13.11.2024 Search for heavy scalar or pseudoscalar states in tt events at CMS | Laurids Jeppe

Overview of the search

- Search for new spin-0 (pseudo)scalars in tt
 final states with full Run 2 dataset (138 fb⁻¹)
- Make use of invariant tt mass, angular and spin correlation observables
- Two analysis channels: dilepton (ll) and lepton+jets (lj)
- Builds upon previous work by CMS: JHEP 04 (2020) 171 (2016 data, 35.9 fb⁻¹)
- Also recent full Run 2 result by ATLAS (JHEP08 (2024) 013)



Signal modeling

- Generic heavy pseudoscalar (A) or scalar (H) coupling solely to top quarks
- Production in gluon fusion via top quark loop



- Same final state as SM $t\bar{t} \rightarrow$ interference \rightarrow peak-dip structure in m_{tt}
- Free parameters: masses, widths, g_A / g_H
- Use NNLO QCD K-factors for normalization

$$\mathcal{L}_A^{\text{int}} = ig_{\text{At}\bar{\text{t}}} \frac{m_{\text{t}}}{v} \bar{\text{t}}\gamma_5 \text{tA}$$

$$\mathcal{L}_{H}^{\mathrm{int}} = -g_{\mathrm{Ht}\bar{\mathrm{t}}} \frac{m_{\mathrm{t}}}{v} \bar{\mathrm{t}}\mathrm{t}\mathrm{H}$$



Modeling: tt bound state effects

- State of the art: non-relativistic QCD (NRQCD)
 - Color-singlet (${}^{1}S_{0}^{[1]}$) attractive
 - \rightarrow Peak below the tt threshold
 - $^{\rm \tiny D}$ Color-octet ($^1{\rm S}_0^{[8]}$ or $^3{\rm S}_1^{[8]}$) repulsive
 - \rightarrow Expected to be small below the $t\bar{t}$ threshold

Lineshape and width not exactly known
 but below experimental resolution





Modeling: tt bound state effects

Use simplified model for MC simulation: η_t

t (PRD 104 (2021) 034023)

- Generic spin-0, color-singlet state η_t
- Couplings to gluons and tops (pseudoscalar)
- Fit mass from NRQCD:

$$m_{\eta_t} - 2m_t = -2 \,\mathrm{GeV} \quad \Rightarrow \quad m_{\eta_t} = 343 \,\mathrm{GeV}$$

- Restrict to $m_{WbWb} \in [337, 349] \, GeV$ to not influence t continuum as predicted by perturbative QCD
- Not available yet: by-event reweighting to NRQCD



Result: very similar signature as low-mass A resonance

13.11.2024 Search for heavy scalar or pseudoscalar states in tt events at CMS | Laurids Jeppe 7 **DESY**.

Lepton+jets channel

- Require exactly one lepton (e/µ), 3 or more jets and 2 or more b tags
- Split into 4 categories: e vs µ and 3 jets vs 4+ jets

(NIM A 736 (2014) 169-178)

- Reconstruct tt system with NeutrinoSolver algorithm:
 - Assign b jets by maximum likelihood
 - Energy correction factor applied for 3 jet events (lost or merged jets) (NIM A 788 (2015) 128-136)



Lepton+jets channel

- Require exactly one lepton (e/µ), 3 or more jets and 2 or more b tags
- Split into 4 categories: e vs µ and 3 jets vs 4+ jets

(NIM A 736 (2014) 169-178)

- Reconstruct tt system with NeutrinoSolver algorithm:
 - Assign b jets by maximum likelihood
 - Energy correction factor applied for 3 jet events (lost or merged jets) (NIM A 788 (2015) 128-136)
- 2D binning in m_{tt} x |cosθ*|
- θ^* : scattering angle of leptonic top quark
 - SM tī: peaks at large cosθ*
 - A/H signal: isotropic \rightarrow flat distribution
 - Sensitive to spin of mediator (but not parity)



Dilepton channel

- Exactly two opposite-sign leptons (e/µ), at least 2 jets, and at least 1 b tag
- Split by lepton flavor: ee, eµ and µµ
- Reject low-m_θ events
 Cut away Z peak & require p_T^{miss} > 40 GeV in ee/μμ



Dilepton channel

- Exactly two opposite-sign leptons (e/µ), at least 2 jets, and at least 1 b tag
- Split by lepton flavor: ee, eµ and µµ
- Reject low-m_{ℓℓ} events
 Cut away Z peak & require p_T^{miss} > 40 GeV in ee/µµ
- Analytic reconstruction of tt system:
- Assumptions: all p_T^{miss} from vv, tops and Ws on-shell
- Assign b jets using likelihood based on $m_{\ell b}$
- Finite detector resolution: repeat reconstruction 100 times with randomly smeared inputs, take weighted average (EPJC 75 (2015) 11, 542; PRD 73 (2006) 054015)



Spin correlation observables

- Both A/H and η_t predict $t\bar{t}$ production in a pure $t\bar{t}$ spin state: 1S_0 or 3P_0 (from A / η_t resp. H)
- Top decays before hadronization \rightarrow transfer spin information to decay products
- Construct spin correlation observables from tops & leptons

Spin correlation observables

- Both A/H and η_t predict $t\bar{t}$ production in a pure $t\bar{t}$ spin state: ¹S₀ or ³P₀ (from A / η_t resp. H)
- Top decays before hadronization \rightarrow transfer spin information to decay products
- Construct spin correlation observables from tops & leptons

- Variable #1: Chel
 - Boost leptons into rest frames of their parent tops \rightarrow Scalar product between directions of flight
 - Straight line with slope sensitive to $t\bar{t}$ spin state ("D")
 - Maximal for ${}^{1}S_{0}$ (from A / η_{t}) separates from SM



Spin correlation observables

- Both A/H and ηt predict tt production in a pure tt spin state: ¹S₀ or ³P₀ (from A / ηt resp. H)
- Top decays before hadronization \rightarrow transfer spin information to decay products
- Construct spin correlation observables from tops & leptons

- Variable #2: Chan
 - Similar as c_{hel}, separating scalars from SM
 - Maximally negative slope for ³P₀ state (from H)
 - Construct similarly from lepton momenta, with sign flip for component parallel to top momentum
- 3 search variables in dilepton: m_{tt} x C_{hel} x C_{han}



14

Background modeling

- Major irreducible background: SM tt
 - Model from NLO MC (Powheg+Pythia)
 - Correct to NNLO QCD and NLO EW from fixed-order predictions by reweighting in 2D bins of m_{tt} and cosθ*

 NNLO QCD: Matrix
 (EPJC 78 (2018) 537)

 NLO EW: Hathor
 (EPJC 51 (2007) 37)

 Normalize to NNLO+NNLL cross section (CPC 185 (2014) 2930)



Background modeling

- Major irreducible background: SM tt
 - Model from NLO MC (Powheg+Pythia)
 - Correct to NNLO QCD and NLO EW from fixed-order predictions by reweighting in 2D bins of m_{tt} and cosθ*

 NNLO QCD: Matrix
 (EPJC 78 (2018) 537)

 NLO EW: Hathor
 (EPJC 51 (2007) 37)

 Normalize to NNLO+NNLL cross section (CPC 185 (2014) 2930)



- Other backgrounds: tW, t channel single-top, rare processes (from MC)
- Z+jets in *l*: from MC with data-driven normalization from Z peak sideband
- QCD+EW processes in l+jets: data-driven shape from sideband with no b tags

Prefit distributions: $\ell + \ge 4$ jets

Differences between data and prediction observed in low m_{tt} bins!



Search for heavy scalar or pseudoscalar states in tt events at CMS | Laurids Jeppe

Prefit distributions: l + 3 jets

Differences between data and prediction observed in low m_{tt} bins!



Search for heavy scalar or pseudoscalar states in tt events at CMS | Laurids Jeppe

Prefit distributions: *ll*

Differences between data and prediction observed in low mtt bins!



Search for heavy scalar or pseudoscalar states in tt events at CMS | Laurids Jeppe

A/H interpretation

- Limits on A or H using only the perturbative QCD+EW background model
- Excess at low m_{tt} visible at low A/H masses stronger for A



13.11.2024

Search for heavy scalar or pseudoscalar states in tt events at CMS | Laurids Jeppe

DESY.

tt bound state?

- Excess is located at low m_{tt} , stronger for pseudoscalar \rightarrow could this be interpreted as $t\bar{t}$ bound state effects?
- Extract cross section using the ηt color-singlet model
 - " "cross section" = difference to perturbative prediction

tt bound state?

- Excess is located at low m_{tt} , stronger for pseudoscalar \rightarrow could this be interpreted as $t\bar{t}$ bound state effects?
- Extract cross section using the ηt color-singlet model
 - " "cross section" = difference to perturbative prediction

 $\sigma(\eta_t) = 7.1 \pm 0.8 \, \mathrm{pb}$

• Agrees with NRQCD prediction: $\sigma(\eta_t)^{\rm pred} = 6.43\,{
m pb}$

(PRD 104 (2021) 034023) (JHEP 09 (2010) 034)

- Word of caution: this model is not a complete description of a tt bound state!
 - missing e.g. color-octet states expected to be small soft initial state gluons – could change color-octet into singlet states etc...

13.11.2024 Search for heavy scalar or pseudoscalar states in tt events at CMS | Laurids Jeppe 22 **DESY**.

Postfit distributions: η_t ($\ell + \ge 4$ jets)

Postfit for η_t model describing the data well



13.11.2024

Search for heavy scalar or pseudoscalar states in tt events at CMS | Laurids Jeppe

Postfit distributions: η_t ($\ell\ell$)

Postfit for η_t model describing the data well



Search for heavy scalar or pseudoscalar states in tt events at CMS | Laurids Jeppe

Postfit distributions: η_t ($\ell\ell$)

Postfit for η_t model describing the data well



Search for heavy scalar or pseudoscalar states in tt events at CMS | Laurids Jeppe



Uncertainties

- Uncertainty on η_t cross section dominated by background modeling
- Leading systematic sources:
 - EW corrections, including SM Top-Higgs Yukawa: $y_t = 1.00^{+0.11}_{-0.12}$ (EPJC 79 (2019) 421)
 - Parton shower scales
 - Missing higher orders
 - PDF
 - Top mass



Checks of the result

- Off-shell effects in tt MC only approximate in Powheg+Pythia (NWA)
 - Check with Powheg bb41 (complete off-shell NLO calculation of pp $\rightarrow b\overline{b}\ell\bar{\ell}\nu\bar{\nu}$)
 - Only available in dilepton for now
 - Redo our extraction with bb41 for the tt+tW prediction in $\ell\ell$ only

Prediction for SM $t\bar{t}$ and tW	Extracted η_t cross section	Uncertainty
b_bbar_41 (POWHEG vRES)	5.9 pb	18%
Default (POWHEG v2)	7.5 pb	13%

NFSY

- Results compatible at ~ 2 SD excess clearly present also with bb41
- Further checks:
 - different generators for SM tt (aMC@NLO+Pythia, Powheg+Herwig)
 - different treatment of NNLO QCD/NLO EW corrections
 - decorrelation of several syst. uncs (e.g. top mass)
- All checks compatible with nominal within uncertainty of result

13.11.2024 Search for heavy scalar or pseudoscalar states in tt events at CMS | Laurids Jeppe

Parity of the excess

- Can we quantify whether the excess is scalar or pseudoscalar?
- Take low-mass A/H resonances as proxies for pure ${}^{1}S_{0}$ and ${}^{3}P_{0}$ tt states
- 2D fit with arbitrary signal strengths

- Data prefers pure ¹S₀ / pseudoscalar
- scalar component compatible with 0 at the level of ~ 2 SD

13.11.2024



A/H limits including η_t

- QCD + η_t describes data well \rightarrow set (BSM) A/H limits
 - $\rightarrow \eta_t$ added as an additional BG process with free-floating normalization



A+H interpretation

- BSM models (e.g. 2HDM) often predict the simultaneous presence of A and H
- Model-independent exclusion contours for both A and H couplings
 - numerical Feldman-Cousins method
- Input for bounds on concrete BSM models



Summary

- Search for new spin-0 (pseudo)scalars in tt final states with full Run 2 dataset
- Dilepton and lepton+jets channels, using m_{tt} , angular and spin observables
- Observed excess in data at low m_{tt} consistent with pseudoscalar
 - Interpretations in terms of a simplified model of a tt
 bound state ηt
 or a generic pseudoscalar A and scalar H
 - Extracted cross section for a parametrized η_t (toy) model (PRD 104 (2021) 034023)
- Set stringent limits on A, H, and A+H with η_t included in the background
- For the future:

An improved non-relativistic QCD calculation of $t\bar{t}$ bound state effects is crucial! Input from theory welcome

Whatever the excess is – it is exciting!

Reference: CMS-PAS-HIG-22-013

13.11.2024 Search for heavy scalar or pseudoscalar states in tt events at CMS | Laurids Jeppe 32 DEST.

Backup

13.11.2024 Search for heavy scalar or pseudoscalar states in tt events at CMS | Laurids Jeppe 33 **DESY**.

Spin density matrix

- Both A/H and η_t predict $t\bar{t}$ production in a pure $t\bar{t}$ spin state: ¹S₀ or ³P₀ (from A / η_t resp. H)
- Encoded in spin density matrix:



- Choose helicity basis $\{\hat{k}, \hat{r}, \hat{n}\}$:
 - \hat{k} : direction of flight of the top quark
 - \hat{r} and \hat{n} : orthogonal to \hat{k}



Definition of c_{hel} and c_{han}

- Start in tt rest frame, boost leptons into rest frames of their parent tops
- Define lepton three-momenta $\hat{\ell}^+$ and $\hat{\ell}^-$ w.r.t $\{\hat{k}, \hat{r}, \hat{n}\}$ basis:
 - \hat{k} : direction of flight of the top quark
 - \hat{r} : orthogonal to \hat{k} in the scattering plane
 - \hat{n} : orthogonal to \hat{k} and \hat{r}

$$c_{\text{hel}} = -(\hat{\ell}^+)_k (\hat{\ell}^-)_k - (\hat{\ell}^+)_r (\hat{\ell}^-)_r - (\hat{\ell}^+)_n (\hat{\ell}^-)_n$$

$$c_{\text{han}} = +(\hat{\ell}^+)_k (\hat{\ell}^-)_k - (\hat{\ell}^+)_r (\hat{\ell}^-)_r - (\hat{\ell}^+)_n (\hat{\ell}^-)_n$$

It can be shown that they follow a straight line with

$$\frac{1}{\sigma} \frac{d\sigma}{dc_{\rm hel}} = \frac{1}{2} \left(1 - D \, c_{\rm hel} \right) \qquad \frac{1}{\sigma} \frac{d\sigma}{dc_{\rm han}} = \frac{1}{2} \left(1 + D^{(k)} \, c_{\rm han} \right) \qquad \text{(JHEP 03 (2024) 099)}$$

13.11.2024Search for heavy scalar or pseudoscalar states in tt events at CMS | Laurids Jeppe35DESY

List of systematic uncertainties

Experimental

- Jet energy corrections split into 11 subsources
- Jet energy resolution
- Unclustered p^T_{miss} (uncorrelated between years)
- Luminosity correlated and decorrelated parts between years
- Pileup
- Trigger efficiencies (separate for ll / lj)
- Electron efficiencies (reco. & ID)
- Muon efficiencies split into syst. and stat.
- B tagging and mistagging efficiencies
- B tagging split into subsources
- L1 ECAL prefiring (where applicable)
- Data-driven EW+QCD BG (*l*+jets) : shape & rate (50%) uncorrelated between channels
- Data-driven Z+jets normalization (ll)

Theory

- Factorization & renormalization scales:
 - $t\bar{t}$, tW, tq, Z+jets; η_t (BG or signal), A/H signal
 - Uncorrelated between processes
 - tt
 : including cross section variation
- Same for initial & final state radiation PS scales
- MC top mass: ±1GeV (interpolated from ±3GeV)
 - Also including cross section variations
- ME-PS matching (h_{damp})
- Underlying event tune
- Color reconnection: 3 different samples
- PDF: PCA performed on final templates from 100 replicas → only leading component considered
- PDF α_s
- Electroweak corrections:
 - SM Higgs-Top Yukawa coupling (1 +0.11 -0.12)
- EW correction scheme (additive v. multiplicative)
- Minor BG cross sections: 15% for tW and tq; 30% for Diboson and tt+X

List of MC generators

Process	QCD order	ME Generator
$t\overline{t}$	NLO	POWHEG V2 (hvq)
tW	NLO	POWHEG V2 (ST_wtch)
Z+jets	NNLO	Powheg v2 (Zj MiNNLO)
t-channel single top	NLO	POWHEG V2 $(ST_tch) + MADSPIN$
s-channel single top	NLO	MG5_AMC@NLO
$t\overline{t}W$	NLO	MG5_AMC@NLO
$t\bar{t}Z$	NLO	MG5_AMC@NLO
WW, WZ & ZZ	LO	Pythia 8.2
A/H signal	LO	MG5_AMC@NLO
$\eta_{ m t} { m signal}$	LO	MG5_AMC@NLO

Data-driven Z+jets normalization

- b jets in Z+jets are known to be badly modeled in MC might lead to wrong normalization after requiring >= 1 btag
- Take normalization from Z peak sideband (R_{in/out} method)
- Use weaker assumption than standard $R_{in/out}$ ("ratio of ratios"): Get $R_{in/out}$ in 0 b tag sideband; take "ratio of ratios" for ≥ 1 and 0 btags from MC

$$\frac{(R_{in/out}^{\geq 1b})_{data}}{(R_{in/out}^{\geq 1b})_{MC}} = \frac{(R_{in/out}^{0b})_{data}}{(R_{in/out}^{0b})_{MC}} \longrightarrow SF = \frac{(N_{out}^{\geq 1b})_{data}}{(N_{out}^{\geq 1b})_{MC}} = \frac{(N_{in}^{\geq 1b})_{data}}{(N_{in}^{\geq 1b})_{MC}} \frac{(R_{in/out}^{0b})_{MC}}{(R_{in/out}^{0b})_{data}}$$

with $N_{data} = N_{data}^{\ell\ell} - 0.5N_{data}^{e\mu}k_{\ell\ell}$, where $k_{ee} = \frac{1}{k_{\mu\mu}} = \sqrt{\frac{N_{data}^{ee}}{N_{data}^{\mu\mu}}}$

13.11.2024 Search for heavy scalar or pseudoscalar states in tt events at CMS | Laurids Jeppe 38 **DESY**.

EW corrections to $t\bar{t}$

- Our EW correction (Hathor) is NLO in EW but LO in QCD
- Ambiguity on how to apply EW corrections to (N)NLO simulation
- Nominal choice: multiplicative

$$\sigma^{\text{rew.}} = \sigma^{\text{LO EW}}_{\text{NLO QCD}} \times \frac{\sigma^{\text{NLO EW}}_{\text{LO QCD}}}{\sigma^{\text{LO EW}}_{\text{LO QCD}}}$$
Alternate choice: additive MadGraph

$$\sigma^{\text{rew.}} = \sigma^{\text{LO EW}}_{\text{NLO QCD}} + \sigma^{\text{NLO EW}}_{\text{LO QCD}} - \sigma^{\text{LO EW}}_{\text{LO QCD}}$$

Lathor

Difference treated as systematic uncertainty

Postfit distributions: A/H

Postfit for A, 365 GeV, 2% width (best fit point)



O DESY.

Search for heavy scalar or pseudoscalar states in tt events at CMS | Laurids Jeppe

Postfit distributions: A/H

Postfit for A, 365 GeV, 2% width (best fit point)



Search for heavy scalar or pseudoscalar states in tt events at CMS | Laurids Jeppe

Postfit distributions: A/H

Postfit for A, 365 GeV, 2% width (best fit point)



Search for heavy scalar or pseudoscalar states in tt events at CMS | Laurids Jeppe

Postfit distributions: η_t

Postfit for η_t



13.11.2024

Search for heavy scalar or pseudoscalar states in tt events at CMS | Laurids Jeppe

Correlation matrix

 Further assess uncertainty modeling through correlations of nuisance parameters



A/H limits including η_t

Limits at different A/H widths



13.11.2024

Search for heavy scalar or pseudoscalar states in tt events at CMS | Laurids Jeppe

A/H limits including η_t

Limits at different A/H widths



13.11.2024

Λ/H limits without η_{t}

Limits at different A/H widths for perturbative QCD background only



47

13.11.2024

Search for heavy scalar or pseudoscalar states in tt events at CMS | Laurids Jeppe

Λ/H limits without η_{t}

Limits at different A/H widths for perturbative QCD background only



13.11.2024

Search for heavy scalar or pseudoscalar states in tt events at CMS | Laurids Jeppe

A+H interpretation



13.11.2024

Search for heavy scalar or pseudoscalar states in tt events at CMS | Laurids Jeppe 49 **DESY**.

Similar full Run 2 ATLAS result: does not see any postfit excess! why?



13.11.2024

- Different channel definitions in ATLAS and CMS
- +jets resolved:
 - ATLAS: 1ℓ, 1b, ≥4 jets
 - CMS: 1ℓ, 2b, 3 jets
 - both: 1ℓ, 2b, ≥4 jets
 → compare pre-fit distributions!
 - Similar prefit excess in data at low m_{tt}!



13.11.2024 Search for heavy scalar or pseudoscalar states in tt events at CMS | Laurids Jeppe 51 **DESY.**

- Different channel definitions in ATLAS and CMS
- l+jets resolved:
 - ATLAS: 1ℓ, 1b, ≥4 jets
 - CMS: 1ℓ, 2b, 3 jets
 - both: 1ℓ, 2b, ≥4 jets
 → compare pre-fit distributions!
 - Similar prefit excess in data at low m_{tt}!
- dilepton: difficult to compare
 - CMS: reconstruct mtt X Chel X Chan
 - ATLAS: no top quark reconstruction instead: m_{ℓℓbb} x Δφ_{ℓℓ}



- Similar prefit excess & expected limits but no postfit excess for ATLAS!
- We are comparing in detail...



SM: tt differential measurements



 \rightarrow good description by theory except for excess in data in threshold region

13.11.2024 Search for heavy scalar or pseudoscalar states in tt events at CMS | Laurids Jeppe 54 **DESY**.

SM: tt spin entanglement

• Measured quantity: D " \approx strength of tt spin correlation" (oversimplified)



13.11.2024 Search for heavy scalar or pseudoscalar states in tt events at CMS | Laurids Jeppe

DESY. 55