Search for Supersymmetry in Boosted Topologies at CMS

SUSY 2017, Mumbai

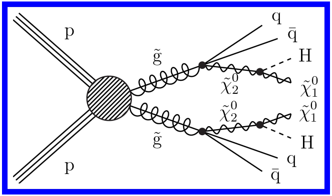
Simon Kurz
on behalf of the CMS Collaboration

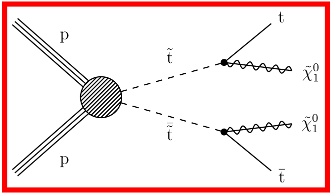
Institut für Experimentalphysik, Universität Hamburg

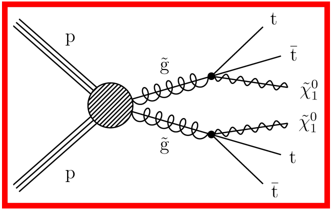
Dec 12, 2017

Scope of the Talk

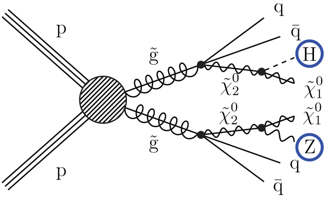
- Focus on two (**new!**) searches that make use of boosted objects
  - **Boosted Higgs**
    CMS-PAS-SUS-17-006
  - **(Boosted) Top quarks**
    CMS-SUS-16-050,
    arXiv:1710.11188
- Magnitude of boost depends on mass splittings
- Similarities:
  - **Strong production** (largest cross-section)
  - **Hadronic decay channels** (jets and missing transverse energy (MET) in final state)







- Several other SUSY searches target decays with **resolved** b-jets
  - See e.g. talk Inclusive search for SUSY in hadronic final states at CMS



→ **Explicitly targets boosted Higgs production**

- Take advantage of $H \to b\bar{b}$ branching fraction of $\approx 58\%$
- High p_T Higgs bosons lead to collimated b-quarks which can be **reconstructed as a single object**

→ Ensured by compressed mass splitting (e.g. $\Delta m(\tilde{g}, \tilde{\chi}_2^0) = 50\,\mathrm{GeV}$, $m(\tilde{\chi}_1^0) = 1\,\mathrm{GeV}$)

- Interpret results for models with H+H and H+Z, i.e. $\mathcal{B}(\tilde{\chi}_2^0 \to H) = 50\%$

1) Baseline Selection

- **Hadronic final states** (also motivated by trigger)
  - Cut on $N_{\text{Jets}} \geq 2$, $H_{\text{T}} > 600\,\text{GeV}$, $H_{\text{T}}^{\text{miss}} > 200\,\text{GeV}$
  - No isolated leptons or tracks
  - Reject if $H_{\text{T}}^{\text{miss}}$ is aligned with leading 4 jets ($\Delta\Phi$)
- **Boosted objects**
  - $E_{\text{T}}^{\text{miss}} > 300\,\text{GeV}$
  - **At least 2 AK8 Higgs candidate jets** with
    - $p_{\text{T}} > 300\,\text{GeV}$
    - Pruned jet mass m_{J} within $50\,\text{GeV}, 250\,\text{GeV}$

2) Signal: Higgs Tagging

- Use events with **1 or 2 Higgs tagged AK8 jets** (N_{Higgs})
  - Pruned jet mass m_{J} within $85\,\text{GeV}, 135\,\text{GeV}$
  - Have a double-b tag (loose WP, $\approx 75\%$ signal eff)
    - link: CMS-PAS-BTV-15-002

Jet Pruning

- Recluster jet
- Remove soft, large angle particles from the jet (often UE, PU)

→ Improves mass resolution

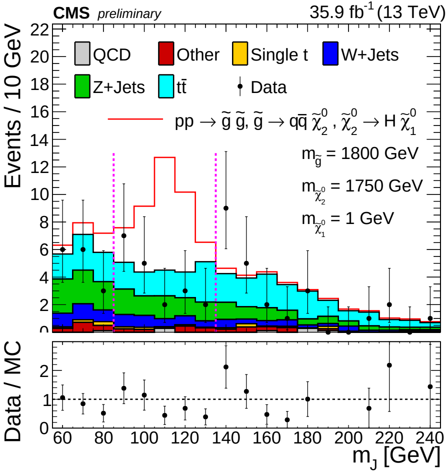
$b\bar{b}$ Tagging

- Based on MVA
- Compatibility of an AK8 jet coming from decay of two b-quarks
- Combines information from secondary vertices, tracks, τ-axes (N-subjettines axes)
- Explicitly doesn't use jet mass m_{J}

→ **More information in backup**

Boosted Higgs: Analysis Design

- Analysis designed as a **counting experiment**
  - $3 \times 2 =$ **6 search regions** ($E_T^{miss} \times N_{Higgs}$)
- **Plot**: Mass m_J of leading jet ($N_{Higgs} = 1, 2$)
  - $\rightarrow$ **Signal** can be seen as a resonance in m_J



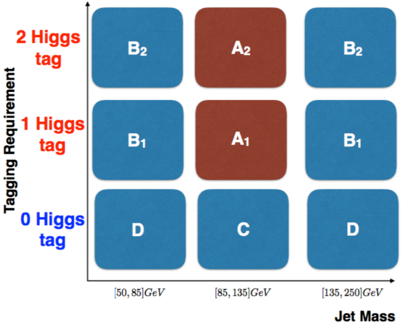
- Estimate the background using an '**ABCD**' method:
  - Use $\boldsymbol{m_J}$ **sideband and un-bb-tagged regions** (background dominated) to extrapolate to **signal region** $\boldsymbol{A_{1,2}}$
  - Measure $\kappa_{1,2}$ in MC to **correct for correlations** of the jet mass and the bb-tagging variables

→ Derive data/MC scalefactors (SFs) in **validation regions** to correct for relative composition of BGs

  - low-$\Delta\Phi$ (QCD)
  - single-lepton ($t\bar{t}$, W+jets)
  - single-photon ($Z(\nu\nu)$+jets)

  → SFs validated in signal sideband regions

- **Validity of ABCD method tested in data** in the validation regions



$$\boxed{A_{1,2} = B_{1,2} \cdot \frac{C}{D} \cdot \kappa_{1,2}}$$

Boosted Higgs: Results

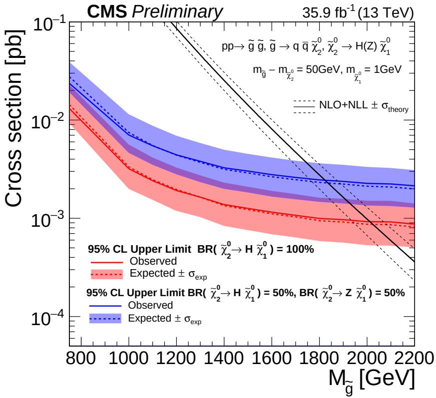
- Observed event yields are in statistical agreement with data-driven prediction of SM backgrounds

$\rightarrow$ **No evidence for SUSY**

N_{Higgs}	$p_{\mathrm{T}}^{\mathrm{miss}}$ (GeV)	κ	Predicted	Observed
1	$300-500$	0.98 ± 0.11	17.68 ± 3.85	15
1	$500-700$	0.86 ± 0.16	3.44 ± 1.47	2
1	> 700	0.86 ± 0.17	0.61 ± 0.45	1
2	$300-500$	0.73 ± 0.14	1.52 ± 0.57	1
2	$500-700$	0.43 ± 0.12	$0.09^{+0.08}_{-0.08}$	0
2	> 700	0.62 ± 0.30	$0.09^{+0.11}_{-0.09}$	0

- Compute upper limits on the gluino pair-production cross section
  - Contributions to sideband regions from signal modeled in likelihood fit
  - **T5HH: exclude $m_{\tilde{g}}$ up to 2010 GeV**
  - **T5ZH: exclude $m_{\tilde{g}}$ up to 1825 GeV**

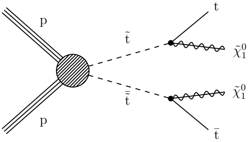
  (with the assumption of $\Delta m(\tilde{g}, \tilde{\chi}_2^0) = 50\,\mathrm{GeV}$, $m(\tilde{\chi}_1^0) = 1\,\mathrm{GeV}$)

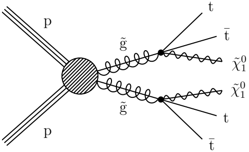


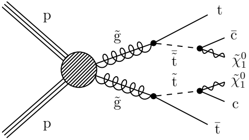
(Boosted) Top: Introduction & Analysis Setup

All-hadronic analysis using top-tagging

- **Jets+E_T^{miss} final states**
  - $N_{Jets} \geq 4,\ E_T^{miss} > 250\,\mathrm{GeV},\ H_T > 300\,\mathrm{GeV},\ N_b \geq 1$
  - Lepton and track vetoes
  - $\Delta\Phi(\text{jet}, E_T^{miss})$-cut on leading three jets
- **Top reconstruction**
  - $N_t \geq 1$ (customized top tagger)
  - $M_{T2} > 200\,\mathrm{GeV}$





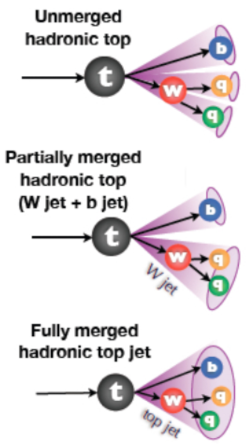


- Targeting large **variety of signal models**
  - $\rightarrow$ **84 search regions** binned in $N_t, N_b, E_T^{miss}, M_{T2}/H_T$
  - Also define **10 aggregate search regions** (simplifies re-interpretation etc.)
- **Customized** top-tagging algorithm
  - Designed for a good efficiency for a **large range of *t*-quark p_T** (boost depends on signal)
  - Uses AK4 and AK8 jets since *t*-jet can be **resolved** or **partially/fully merged**

1. Tag **fully merged** tops
   - **Standard** PUPPI AK8 **boosted top** (loose working point)

2. Tag **partially merged** tops (merged W-jet)
   - **Standard** PUPPI AK8 **boosted W** (loose working point)
   - **Combine with** nearby AK4 jet
   - Require mass to be consistent with $m(t)$ and ratio consistent with $m(W)/m(t)$

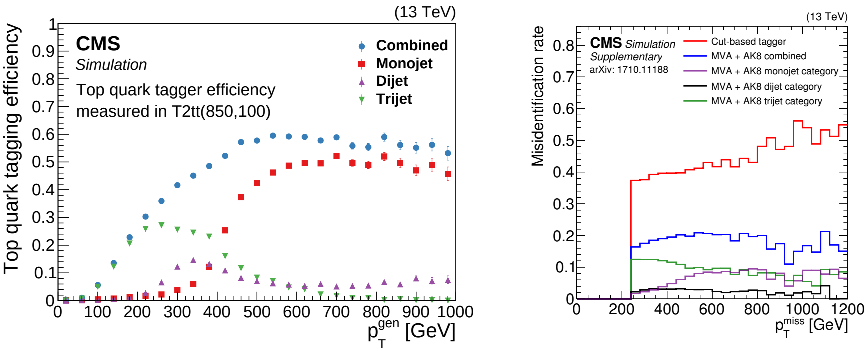
3. Tag **resolved** tops (**New!**)
   - Use **all combinations** of 3 AK4 jets (no more than 1 b-jet)
   - **Train MVA** on jet properties (p_{T}, mass of tri-jet and di-jet system, angular separation, CSV, QGL, etc.)
   - Optimized cut on MVA based on full limit-setting procedure



PUPPI (pileup per particle identification)

- Each particle is weighted by the probability to originate from the primary vertex before jet clustering

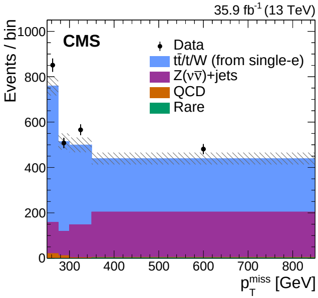
(Boosted) Top: Top Tagger (Performance)

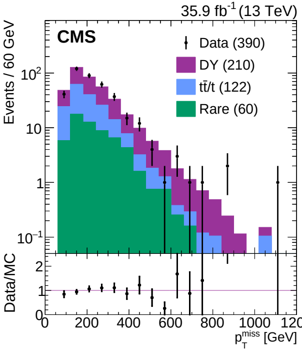


- **Left**: Performance of "Tight" working point in signal MC (T2tt)
  - **High efficiency** for a large range of t-quark p_T
  - Breakdown in the categories reveals important contributions
- **Right**: Misidentification rate of the top tagger in MC ($Z \to \nu\nu$)
  - Almost constant with an average of $\approx 20\%$
  - **Huge improvement** compared to previous (cut-based) version of tagger

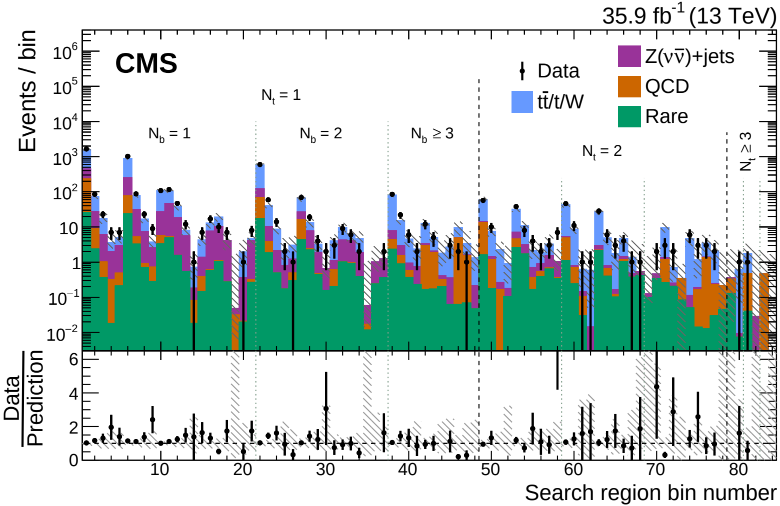
- **Most important backgrounds**: lost lepton and hadronic tau, as well as Z invisible
  1. $t\bar{t}$, W +jets: Data-driven approach using **translation factors**
     - Derive $\mathrm{TF} = N^{\mathrm{MC}}_{\mathrm{SR}} / N^{\mathrm{MC}}_{\mathrm{CR}}$ in MC (data/MC SFs applied)
     - Then estimate using data CR: $N_{\mathrm{SR}} = \mathrm{TF} \cdot N_{\mathrm{CR}}$
     - Use data sideband to validate approach **[top plot]**
  2. $Z(\nu\nu)$ +jets: Central value from **MC with data/MC corrections**
     - Derive SF(N_{Jets}) in "loose" ($DY \to \mu\mu$) CR and apply on ($Z \to \nu\nu$) MC **[bottom plot, after application of SF]**
     - Then adjust normalization in "tight" CR
     - Derive shape uncertainty based on "loose" CR after all corrections have been applied

→ Details on QCD, ttZ background estimation in backup





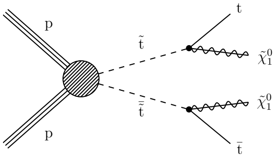
No statistically significant deviation between data and SM background observed

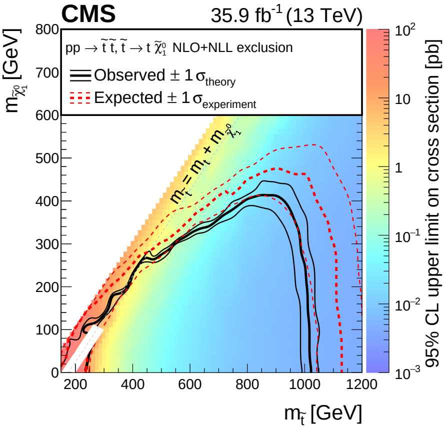


→ Results of aggregated Search Regions in backup!

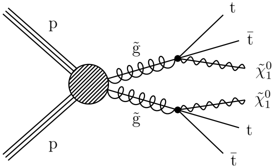
(Boosted) Top: Exclusion Limits **[more in backup]**

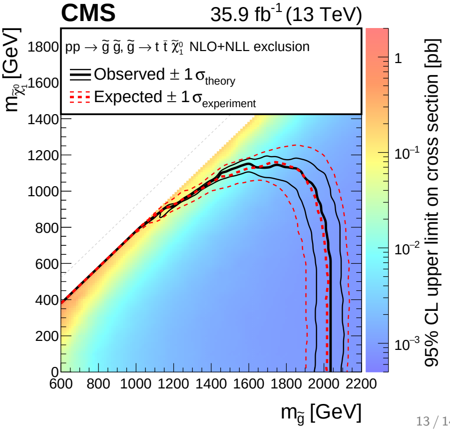
T2tt





T1tttt

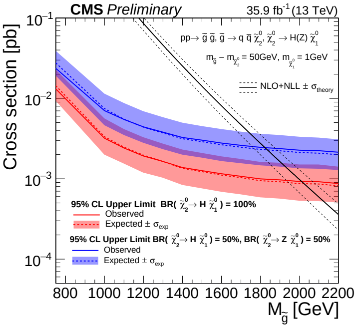


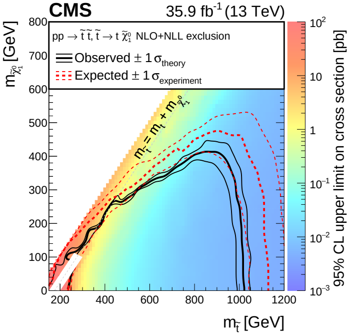


- Highlighted results from two (brand-new!) analyses with boosted final states with $35.9\,\mathrm{fb}^{-1}$ data collected by CMS
  - Both searches for Supersymmetry are performed in jets+MET final state, plus
    1. **boosted Higgs** (CMS-PAS-SUS-17-006, paper will be available soon)
    2. **(boosted) top quark** (link: arXiv:1710.11188)

  $\rightarrow$ No evidence for physics beyond the Standard Model found

- **Additional material for interpretation: link (boosted top)**

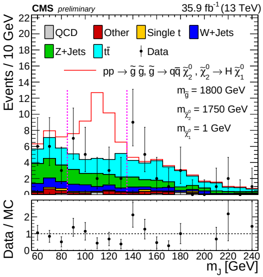




Backup

- Prediction can be done in data using the signal trigger:

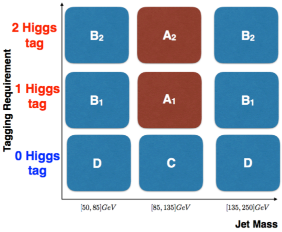
$$\boxed{A_{1,2} = B_{1,2} \cdot \frac{C}{D} \cdot \kappa_{1,2}}$$



- Based on assumption that the jet mass and the bb-tagging variables are un-correlated
  - But: **ratios** $\left(\frac{B_1}{D}, \frac{A_1}{C}, \ldots\right)$ **are sample dependent**

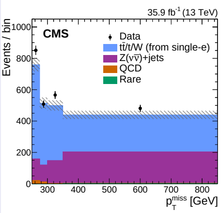
  → Derive correction for the effect:
  $$\kappa_{1,2} = A^{MC}_{1,2} / \left(B^{MC}_{1,2} \cdot \frac{C^{MC}}{D^{MC}}\right)$$
  - $\kappa_{1,2}$ calculated inclusively in all BG MC samples (but independently for all 6 SRs)

- **Derive data/MC SFs in validation regions to correct for relative composition of BGs**
  - Loosen cut on E_T^{miss} or $\Delta\Phi$ in case of low statistics
  - Generally no E_T^{miss} dependence observed so calculated inclusively
  - Single lepton validation region: derive SF as a function of E_T^{miss} and assign E_T^{miss} shape uncertainty in case of low statistics

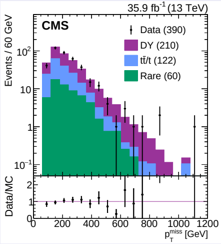


1. ***$t\bar{t}$, W +jets***: Data-driven approach using **translation factors**
   - Derive $\mathrm{TF} = N_{\mathrm{SR}}^{\mathrm{MC}} / N_{\mathrm{CR}}^{\mathrm{MC}}$ in MC (data/MC SFs applied)
   - Then estimate using data CR: $N_{\mathrm{SR}} = \mathrm{TF} \cdot N_{\mathrm{CR}}$
   - Use data sideband to validate approach
2. ***$Z(\nu\nu)$ +jets***: Central value from **MC with data/MC corrections**
   - Derive SF(N_{Jets}) in "loose" ($DY \to \mu\mu$) CR and apply on ($Z \to \nu\nu$) MC
   - Then adjust normalization based on "tight" DY CR
   - Derive shape uncertainty based on "loose" CR after all corrections have been applied
3. **QCD**: Data-driven approach using **translation factors**
   - Measure $\mathrm{TF} = N_{\mathrm{QCD}}^{\Delta\Phi} / N_{\mathrm{QCD}}^{\overline{\Delta\Phi}}$ in data in low $E_{\mathrm{T}}^{\mathrm{miss}}$ sideband region
   - Use MC to extrapolate to SR
   - Derive systematic uncertainty based on MC closure test
4. ***ttZ***: **MC**, validated in three-lepton CR

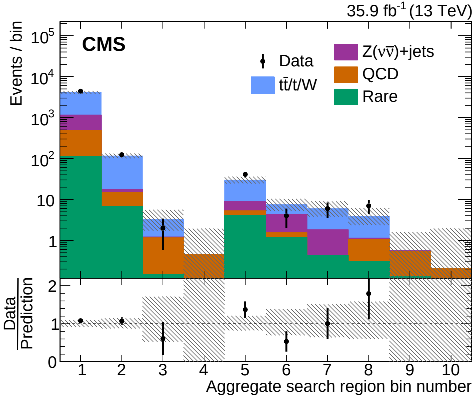
Validation of $t\bar{t}$, W +jets estimation in data sideband



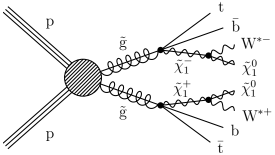
Data/MC in "loose" DY CR after application of SF(N_{Jets})



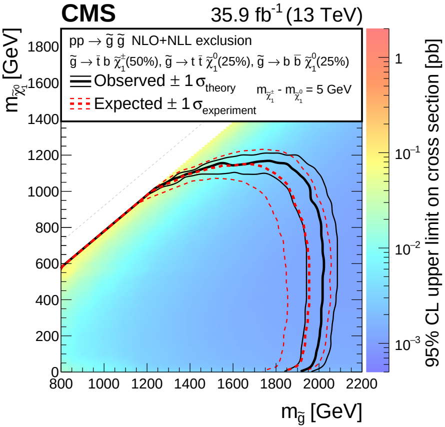
(Boosted) Top: Results (Aggregated Search Regions)



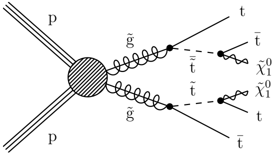
Region	N_t	N_b	m_{T2} [GeV]	p_T^{miss} [GeV]	Motivation
1	≥1	≥1	≥200	≥250	Events satisfying selection criteria
2	≥2	≥2	≥200	≥250	Events with $N_t \geq 2$ and $N_b \geq 2$
3	≥3	≥1	≥200	≥250	Events with $N_t \geq 3$ and $N_b \geq 1$
4	≥3	≥3	≥200	≥250	T5tttt; small $\Delta m(\tilde{g}, \tilde{\chi}_1^0)$ and $m_{\tilde{\chi}_1^0} < m_t$
5	≥2	≥1	≥200	≥400	T2tt; small $\Delta m(\tilde{t}, \tilde{\chi}_1^0)$
6	≥1	≥2	≥600	≥400	T2tt; large $\Delta m(\tilde{t}, \tilde{\chi}_1^0)$
Region	N_t	N_b	H_T [GeV]	p_T^{miss} [GeV]	Motivation
7	≥1	≥2	≥1400	≥500	T1ttbb & T5ttcc; large $\Delta m(\tilde{g}, \tilde{\chi}_1^0)$
8	≥2	≥3	≥600	≥350	T1tttt; small $\Delta m(\tilde{g}, \tilde{\chi}_1^0)$
9	≥2	≥3	≥300	≥500	T1/T5tttt & T1ttbb; intermediate $\Delta m(\tilde{g}, \tilde{\chi}_1^0)$
10	≥2	≥3	≥1300	≥500	T1/T5tttt; large $\Delta m(\tilde{g}, \tilde{\chi}_1^0)$

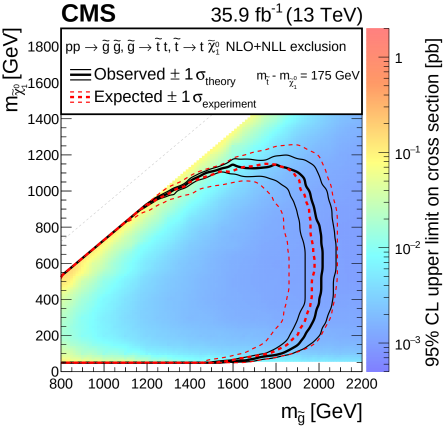


T1ttbb



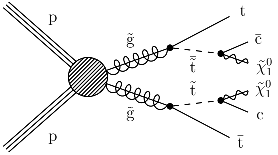
- Gluino decays via an off-shell top or bottom squark
  - $\tilde{g} \to t\bar{t}\tilde{\chi}_1^0(25\%)$
  - $\tilde{g} \to b\bar{t}\tilde{\chi}_1^+$(or charge conjugate,50%)
  - $\tilde{g} \to b\bar{b}\tilde{\chi}_1^0(25\%)$
- $\tilde{\chi}_1^+$ decays to LSP and off-shell W
- Provides sensitivity to mixed states of top and bottom quarks

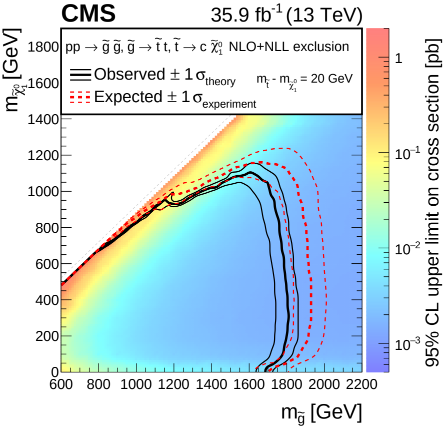




T5tttt

- Gluino decays via on-shell top squark
  - $\Delta m(\bar{t}, \tilde{\chi}^0_1) = 175\,\mathrm{GeV}$
- Provides sensitivity to a region that is difficult to probe with the T2tt model
  - Signal similar to $t\bar{t}$ events
- Area at the bottom is excluded due to high signal contamination
  - Statistical treatment of signal contamination becomes unreliable

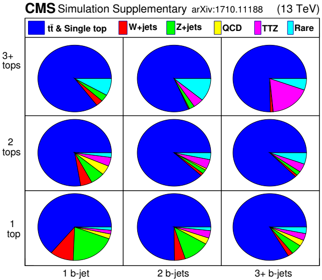


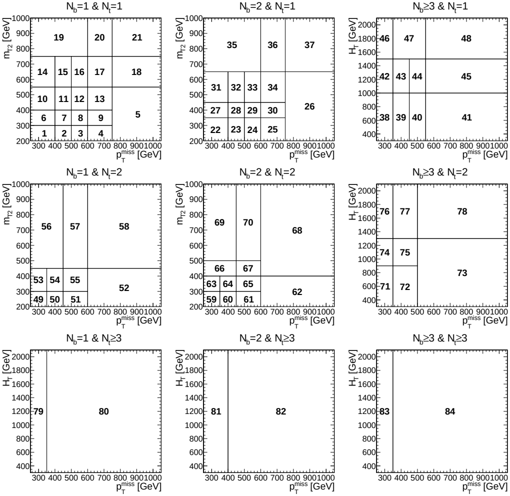


T5ttcc

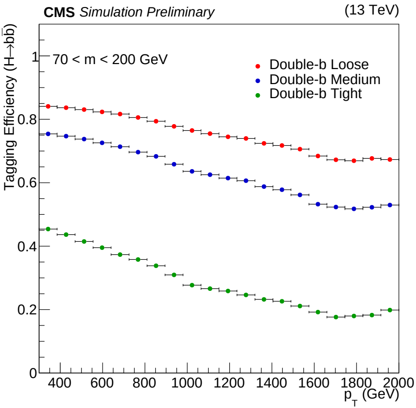
- Top squark decays to a charm quark and the LSP.
  - $\Delta m(\bar{t}, \tilde{\chi}_1^0) = 20\,\mathrm{GeV}$
- Provides sensitivity to scenarios in which the top squark is kinematically unable to decay to an on-shell top quark
  - Dominant decay mode of top squark in this case

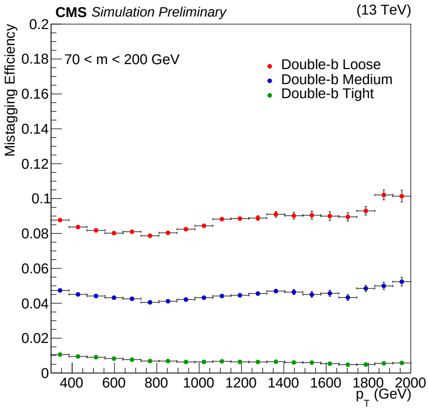
(Boosted) Top: Search Binning



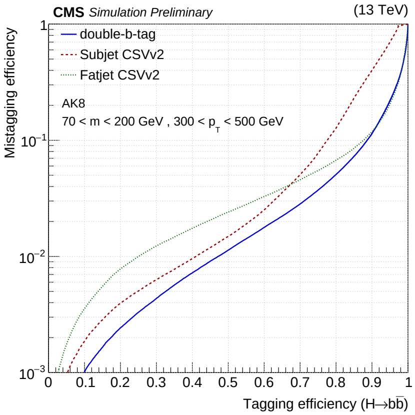


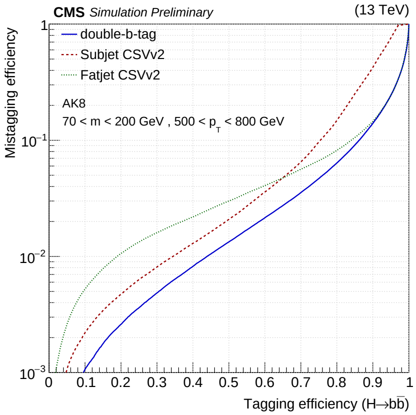
CMS-PAS-BTV-15-002





CMS-PAS-BTV-15-002





(Boosted) Higgs: Performance of $H \to b\bar{b}$ tagger

CMS-PAS-BTV-15-002