Searches for top squarks in compressed scenarios with the CMS experiment

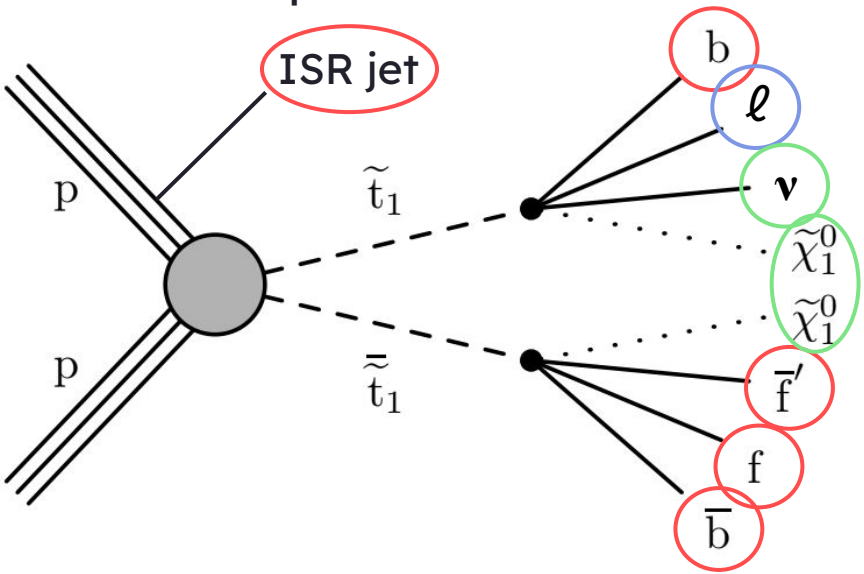
Diogo de Bastos
on behalf of the CMS collaboration

The XXIX International Conference on Supersymmetry and Unification of Fundamental Interactions (SUSY 2022) - 29/06/2022

TÉCNICO LISBOA
FCT Fundação para a Ciência e a Tecnologia
LIP LABORATÓRIO DE INSTRUMENTAÇÃO E FÍSICA EXPERIMENTAL DE PARTÍCULAS
CMS

Characteristics of the stop compressed scenario

Ex: stop 4-body decay in 1 lepton final states



$$\Delta m = m(\tilde{t}_1) - m(\tilde{\chi}_1^0) < m(W)$$

Compressed scenario:
Δm(stop,neutralino) < 80 GeV

LSP can be WIMP Dark Matter that was depleted in the early universe through co-annihilation process to match the observed DM density

- ISR jet to boost the sparticle pair
- Favors soft decay products that can be jets or leptons
- High p_T^{miss} due to Neutralino escaping undetected

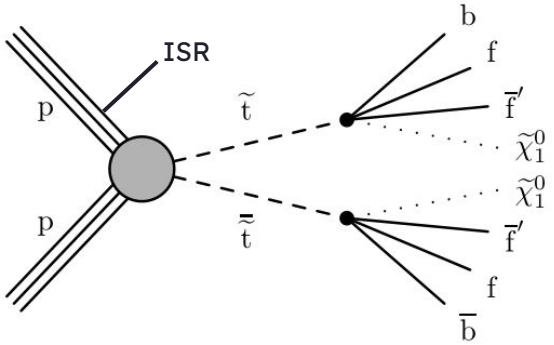
Advances in compressed stop searches at CMS

In this talk we will cover:

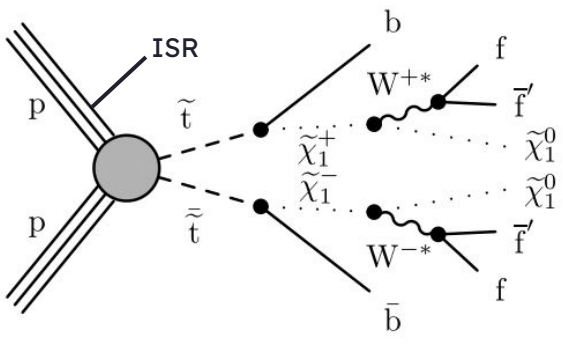
- CMS-SUS-19-010: *Search for top squark production in fully-hadronic final states in proton-proton collisions at √s=13 TeV*
  PRD 104, 052001 (2021) | end of 2021
  Strategy: **Soft b-tag algorithm**
- CMS-SUS-18-004: *Search for SUSY in final states with two or three soft leptons and missing transverse momentum in proton-proton collisions at √s=13 TeV*
  JHEP04(2022)091 | Spring 2022
  Strategy: **Opposite sign lepton pair and soft b-tagged jets**
- CMS-SUS-21-003: *Search for top squarks decaying via the four-body mode in single-lepton final states from Run 2 of the LHC*
  CMS-PAS-SUS-21-003 | Preliminary public results Summer 2022
  Strategy: **Train 1 BDT per Δm**

All-hadronic stop search

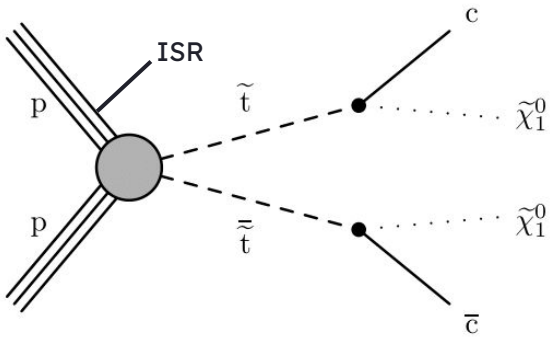
4-body decay



Chargino mediated



2-body decay

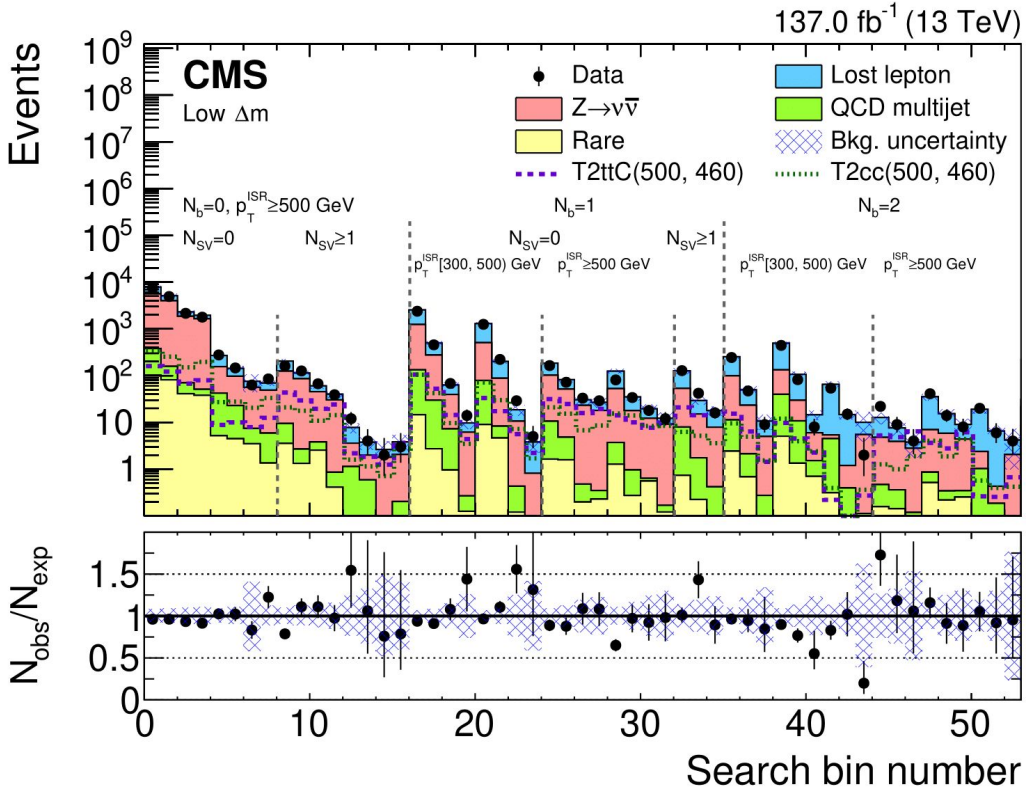


- 4bd: Effective four-point interaction
- Chg: Virtual W decays
- 2bd: Loop-induced flavour-changing neutral current charm production

Using a **cut-and-count** approach, this search takes advantage of b-/c-tagging and soft b-tagging algorithms in all-hadronic final states

PRD 104, 052001 (2021)

All-hadronic Signal Region

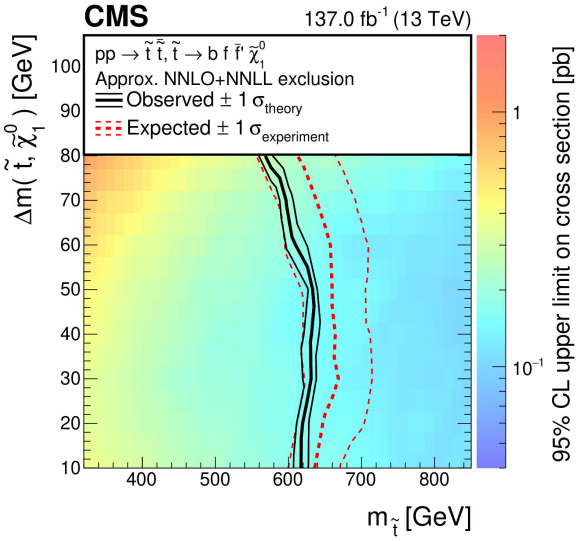


- Online trigger on p_T^{miss}
- H_T > 300 GeV
- p_T^{miss} > 250 GeV
- p_T^{ISR} > 200 GeV
- Veto leptons
- Bins only accessible through soft b-tag (N_{SV})
- SM Backgrounds estimated and validated from control regions

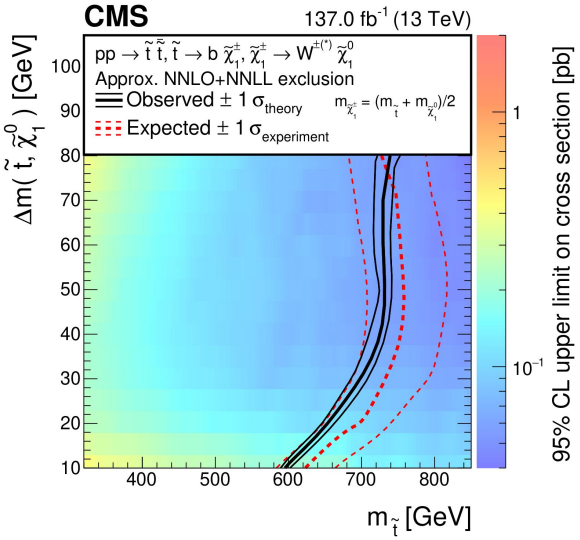
PRD 104, 052001 (2021)

All-hadronic Exclusion Limits

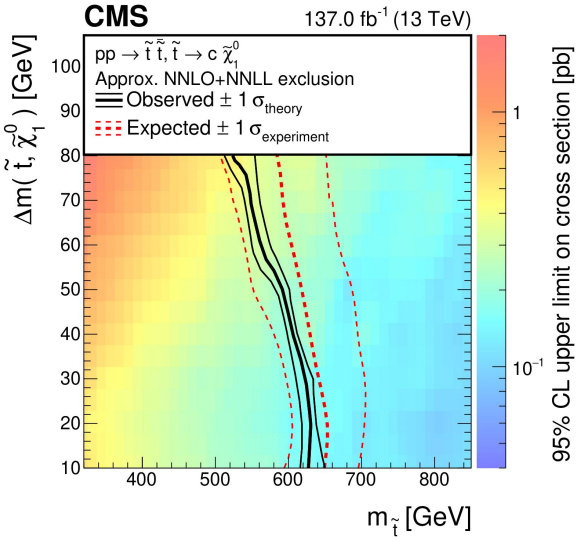
4bd



Chg



2bd



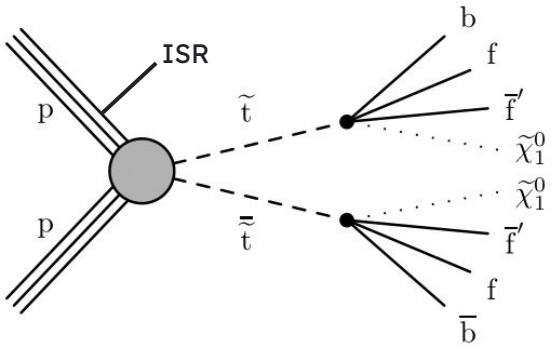
Excluded top squark masses up to:

- 640 GeV for 4bd model
- 740 GeV for Chg model
- 630 GeV for 2bd model.

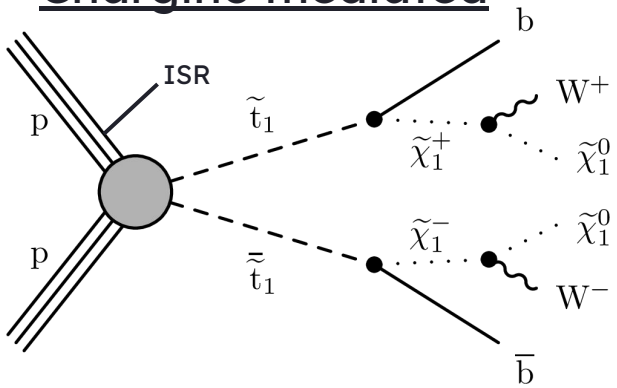
PRD 104, 052001 (2021)

Two-leptons stop search

4-body decay



Chargino mediated

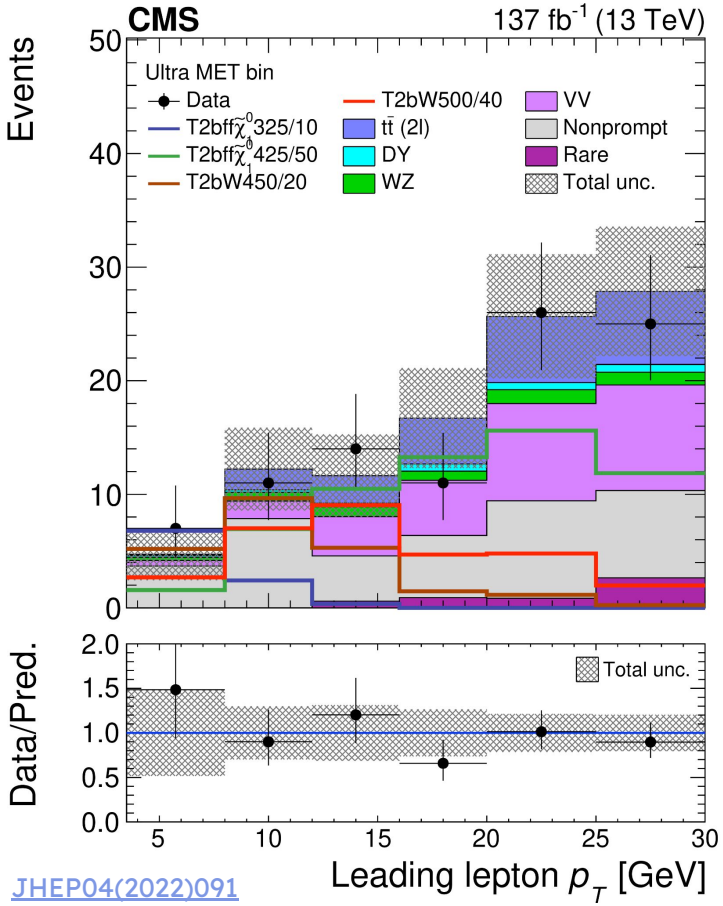


- 4bd: Effective four-point interaction
- Chg: Virtual W decay

Using a **cut-and-count** approach, this search looks for opposite sign lepton pairs final states and soft be tagged jets to probe for 4-body decay and Chargino mediated signal models

JHEP04(2022)091

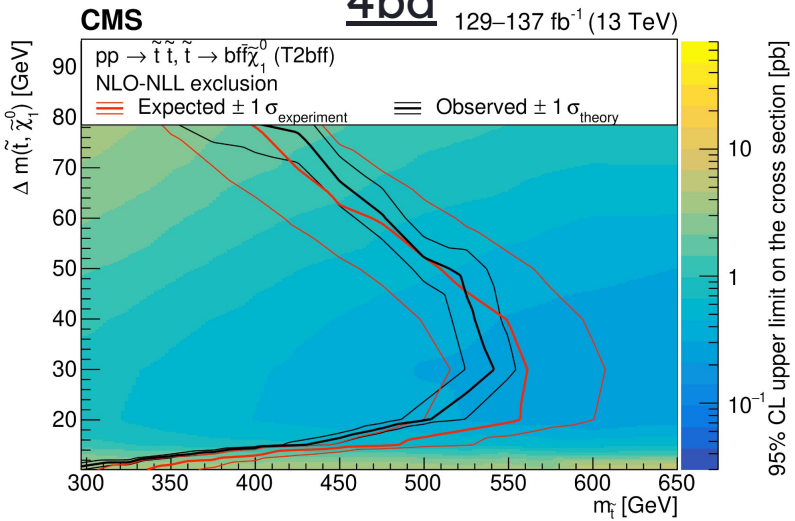
Two-leptons Signal Region

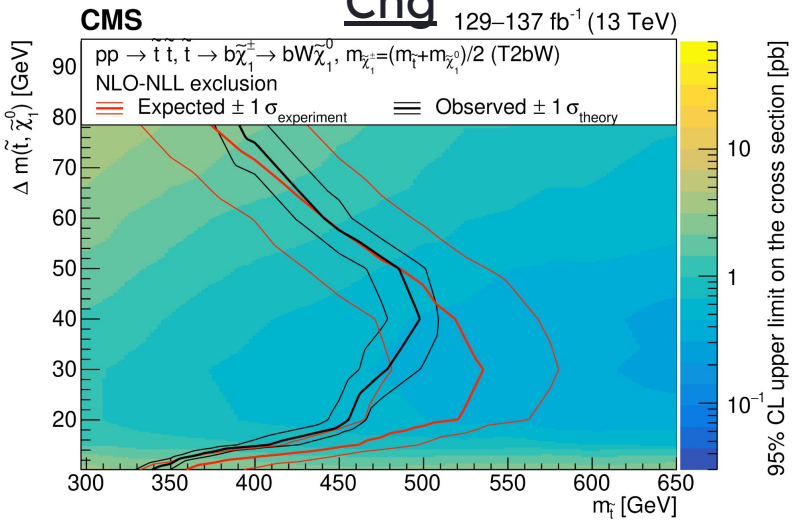


- Online trigger on p_T^{miss} or di-lep+p_T^{miss}
- Search region is first split in p_T^{miss} bins
- Soft leptons: 5(3.5) < $p_T(\ell)$ < 30 GeV for e(μ)
- H_T > 100 GeV
- 0.66 < p_T^{miss}/H_T < 1.4
- “tight lepton veto” criteria on the leading jet → realized only by ISR jet
- Veto b-tagged jet with p_T > 25 GeV
- Prompt lepton backgrounds estimated from CRs
- Non-prompt lepton backgrounds estimated from “tight-to-loose” method

JHEP04(2022)091

Two-leptons Exclusion Limits





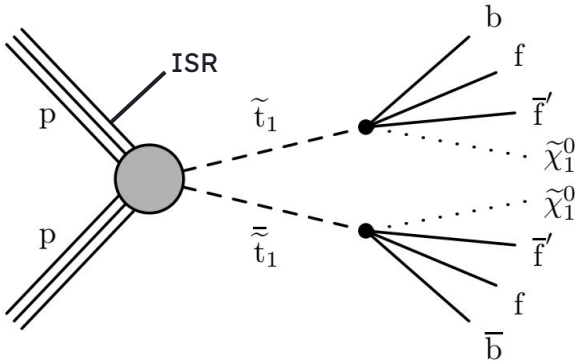
Excluded top squark masses up to:

- 540 GeV for 4bd model
- 480 GeV for Chg model

This search has access to both models but is not optimized for. Used to confirm the results from other specified searches

JHEP04(2022)091

Single-lepton stop search



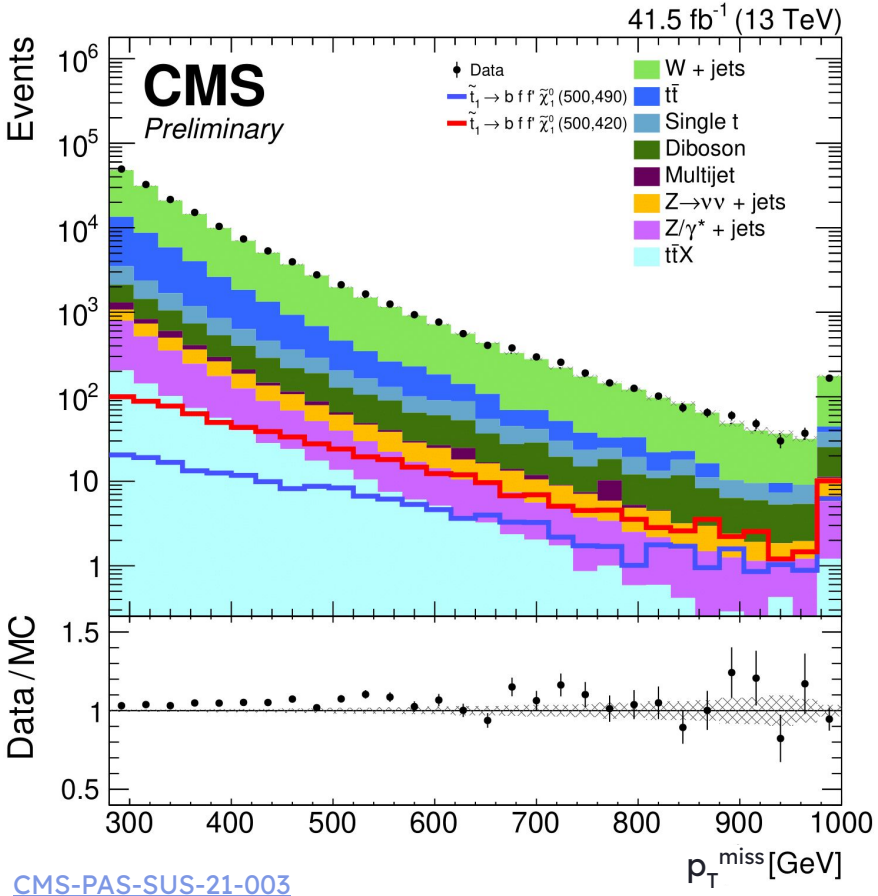
- 4bd: Effective four-point interaction

Using a multivariate approach, this search builds the SRs in 2 steps:

- Preselection region: reduce background while preserving signal
- Train a Boosted Decision Tree (BDT) per Δm region (10 to 80 GeV, in 10 GeV steps) and optimize the cut on the BDT output

CMS-PAS-SUS-21-003

Single-lepton Preselection



- Online trigger on p_T^{miss}
- Exactly 1 lepton
- Soft leptons:
  $p_T(\ell) > 5(3.5)$ GeV for e(μ)
- For $\Delta m \leq 60$ GeV: $p_T(\ell) < 30$ GeV
- $H_T > 200$ GeV
- $p_T^{miss} > 280$ GeV
- $p_T^{ISR} > 110$ GeV

CMS-PAS-SUS-21-003

BDT for the determination of the Signal Regions

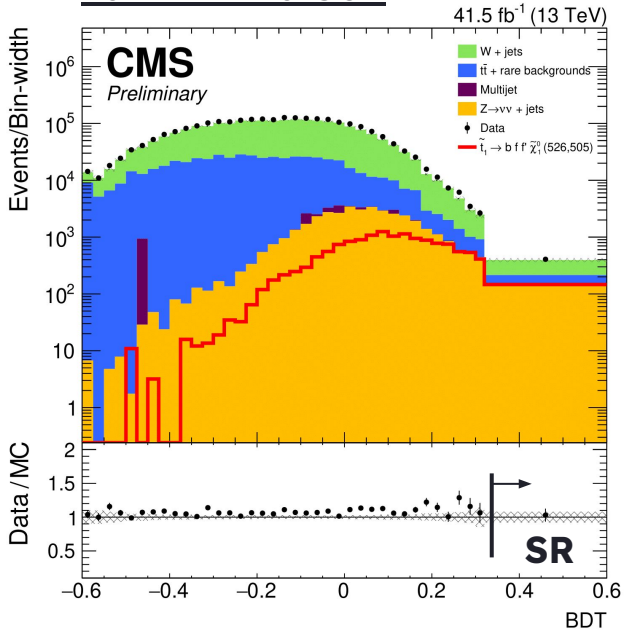
1. Train a BDT per Δm (10 to 80 GeV, in 10 GeV steps):
   - Signal: All the Signal Points (SPs) with the same Δm:
     - SPs with **different Δm** have **different** kinematic distributions
     - SPs with the **same Δm** have **similar** kinematic characteristics
   - Background: WJets+TTbar normalized to XS
2. 12 Discriminant variables: p_T(lep), η(lep), chg(lep), p_T^{miss}, M_T, N_{jets}, $p_T(jet_1)$, $p_T(jet_b)$, H_T, N_{jets}(b loose), ΔR(lep, jet_b) & Disc(jet_b)
3. Check for overtraining
4. Measure signal/background separation as a function of the BDT cut: Selection = Preselection && BDT > **x**
5. The optimal **x** is going to be the cut that **maximizes** this separation
   → Thus, **Signal Region** has been determined

repeat once per Δm, per year

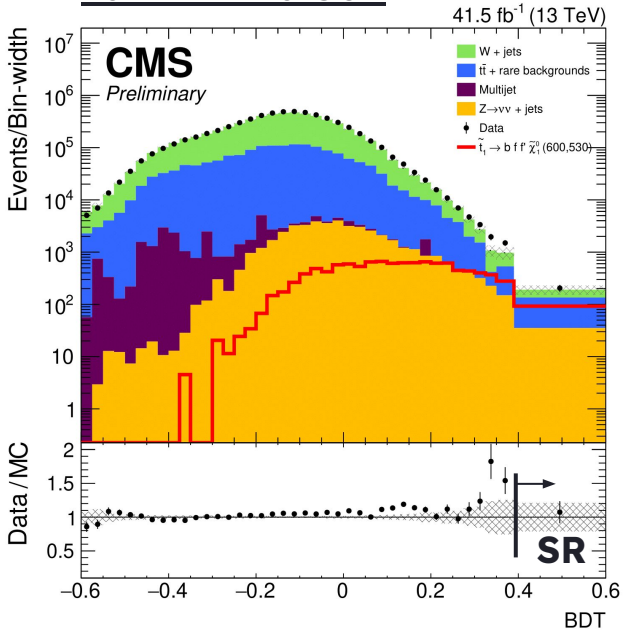
CMS-PAS-SUS-21-003

Single-lepton SRs determined by the BDT output

2017 Δm=20 GeV



2017 Δm=70 GeV

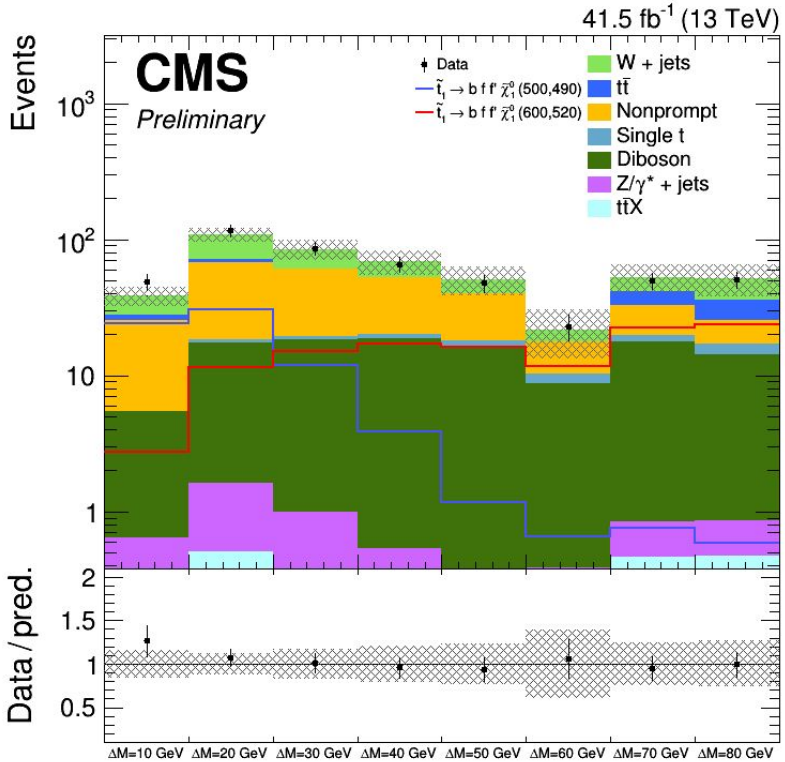


MC estimated backgrounds

BDTs transform the background/signal differences across 12 discriminant variables into a 1-D variable where a cut is set to determine the SR of the respective Δm search region.

CMS-PAS-SUS-21-003

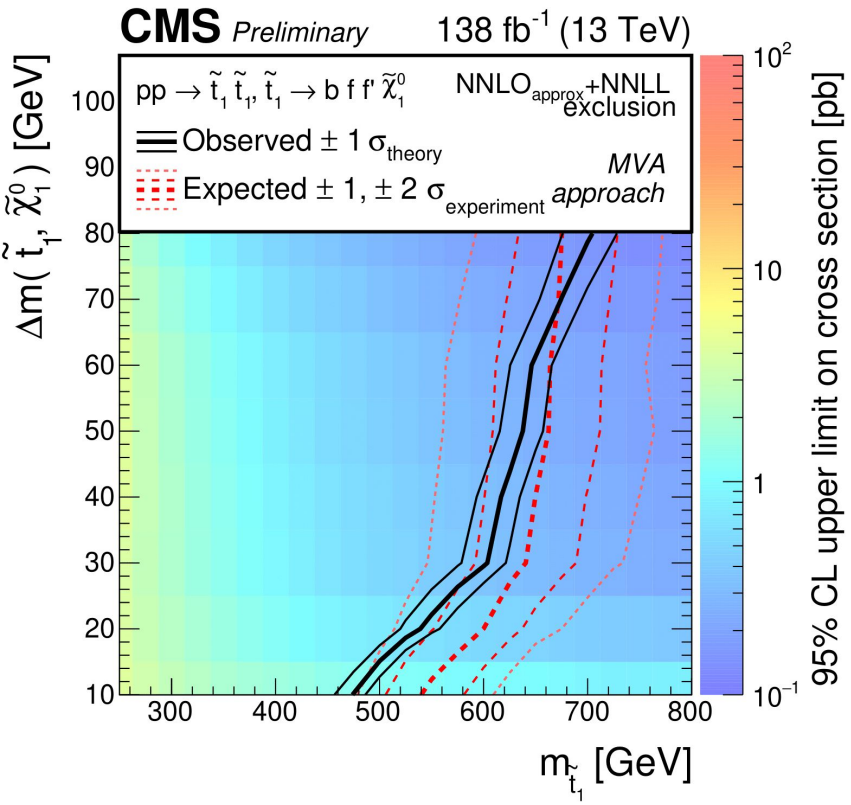
Single-lepton Signal Region



- Each signal region determined by Preselection + $BDT_{cut}(\Delta m)$
- WJets and TTbar prompt backgrounds predicted by data-driven method
- Non-prompt lepton backgrounds estimated from “tight-to-loose” method

CMS-PAS-SUS-21-003

Single-lepton Exclusion Limits



Excluded top squark masses up to:

- 480 GeV at Δm=10 GeV
- 700 GeV at Δm=80 GeV

Local significance of 2.5 standard deviations only at Δm=10 GeV region

CMS-PAS-SUS-21-003

Summary

- Presented the latest results on the searches for stop in compressed scenarios with the CMS experiment
- Covered 0-lepton (all-hadronic), 1-lepton and 2-leptons final states of the 4-body decay of stop in this regimen
- Showcased the different search strategies: soft b-tagging, cut-and-count and multivariate approaches
- No hint of SUSY in the compressed scenario, *yet*
- **Run3:** increase in luminosity, improvement of reconstruction algorithms for soft physics objects and the use of multivariate techniques → will bring interesting searches opportunities…

…Let's keep looking!

Backup

Soft b-tagging

Based on the presence of a Secondary Vertex (SV) reconstructed with inclusive vertex finder (IVF)

- The distance in the transverse plane between the SV and PV < 3 cm
- The significance of the distance between the SV and the PV is > 4
- The pointing angle, defined as $cos((\overrightarrow{PV, SV}, \overrightarrow{p}_{SV}))$ > 0.98, $\overrightarrow{p}_{SV}$ is the total four-momentum of the tracks associated to the SV
- The number of tracks associated to the SV ≥ 3
- The p_T of SV < 20 GeV
- The distance to any jet with p_T > 20 GeV ≥ 0.4 to achieve the orthogonality to the jets and b-tagged jets

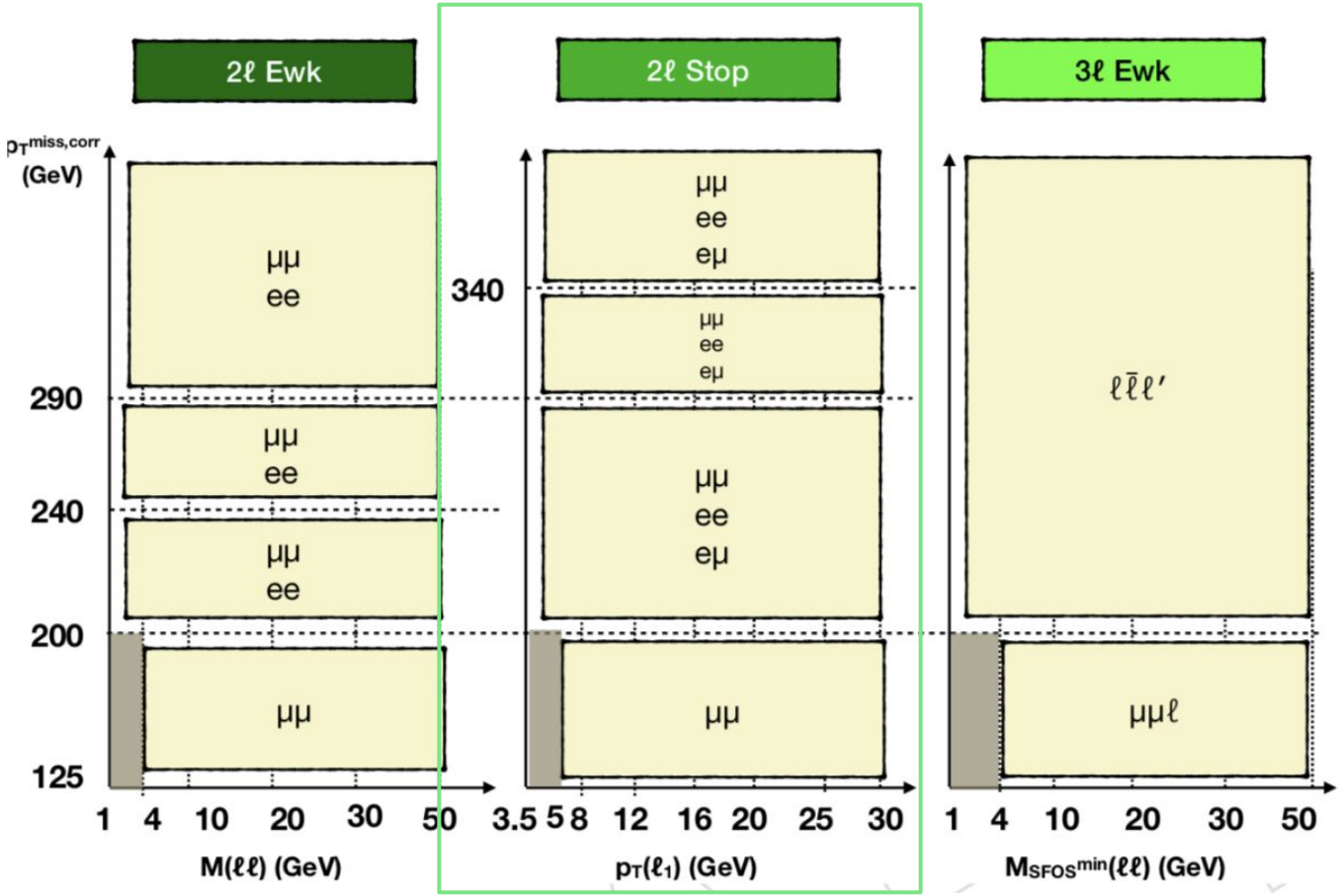
All-hadronic search bin definitions

N_j	N_b	N_{SV}	m_T^b [GeV]	p_T^{ISR} [GeV]	p_T^b [GeV]	p_T^{miss} [GeV]	Bin number
2–5	0	0	—	>500	—	[450, 550, 650, 750, ∞]	0–3
≥6	0	0	—	>500	—	[450, 550, 650, 750, ∞]	4–7
2–5	0	≥1	—	>500	—	[450, 550, 650, 750, ∞]	8–11
≥6	0	≥1	—	>500	—	[450, 550, 650, 750, ∞]	12–15
≥2	1	0	<175	300–500	20–40	[300, 400, 500, 600, ∞]	16–19
≥2	1	0	<175	300–500	40–70	[300, 400, 500, 600, ∞]	20–23
≥2	1	0	<175	>500	20–40	[450, 550, 650, 750, ∞]	24–27
≥2	1	0	<175	>500	40–70	[450, 550, 650, 750, ∞]	28–31
≥2	1	≥1	<175	>300	20–40	[300, 400, 500, ∞]	32–34
≥2	≥2	—	<175	300–500	40–80	[300, 400, 500, ∞]	35–37
≥2	≥2	—	<175	300–500	80–140	[300, 400, 500, ∞]	38–40
≥7	≥2	—	<175	300–500	>140	[300, 400, 500, ∞]	41–43
≥2	≥2	—	<175	>500	40–80	[450, 550, 650, ∞]	44–46
≥2	≥2	—	<175	>500	80–140	[450, 550, 650, ∞]	47–49
≥7	≥2	—	<175	>300	>140	[450, 550, 650, ∞]	50–52

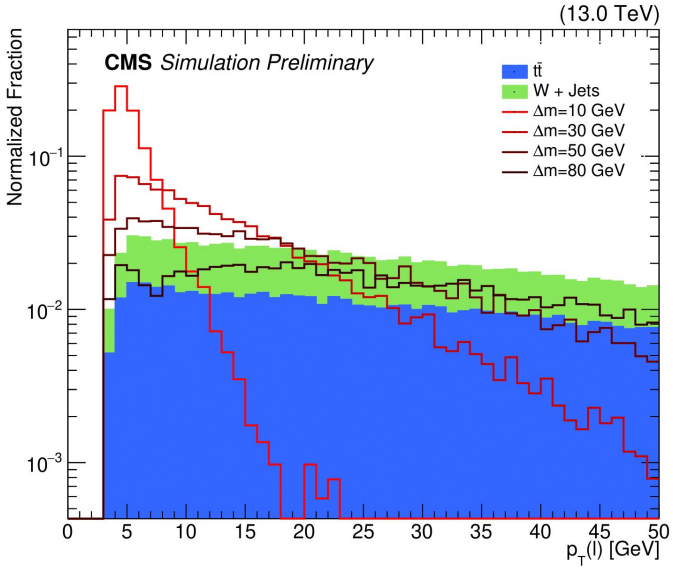
Two-leptons p_T^{miss} SR split

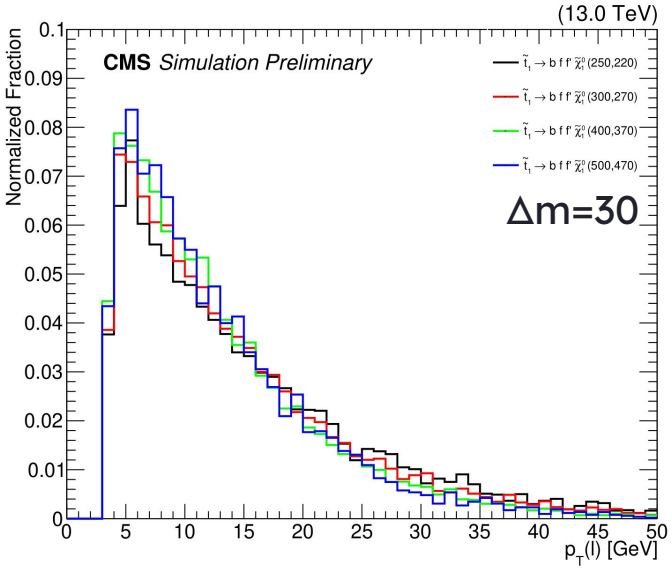
Search region	Low-MET		Med-MET	High-MET	Ultra-MET
	Raw p_T^{miss}	p_T^{miss}	p_T^{miss}	p_T^{miss}	p_T^{miss}
2ℓ-Ewk	> 125	$(125, 200]$	$(200, 240]$	$(240, 290]$	> 290
2ℓ-Stop	> 125	$(125, 200]$	$(200, 290]$	$(290, 340]$	> 340
3ℓ-Ewk	> 125	$(125, 200]$		> 200	

Two-leptons final states



BDT Training Strategy





1 - Signal points with **different Δm** have **different** kinematic distributions

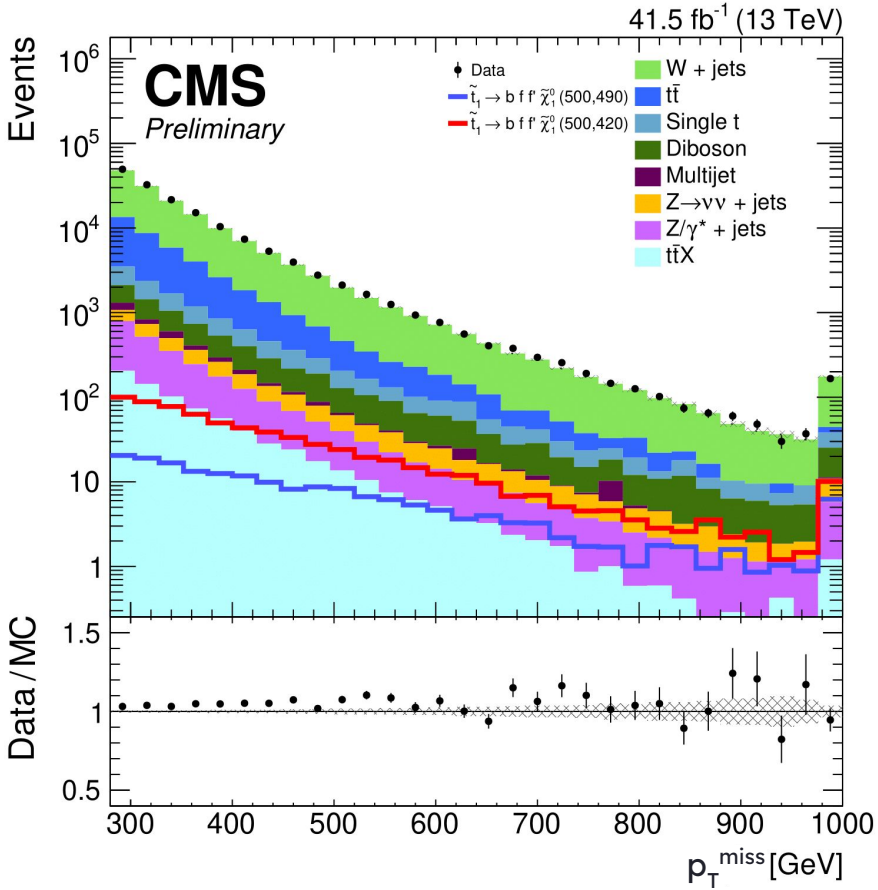
2 - Signal points with the **same Δm** have **similar** kinematic characteristics

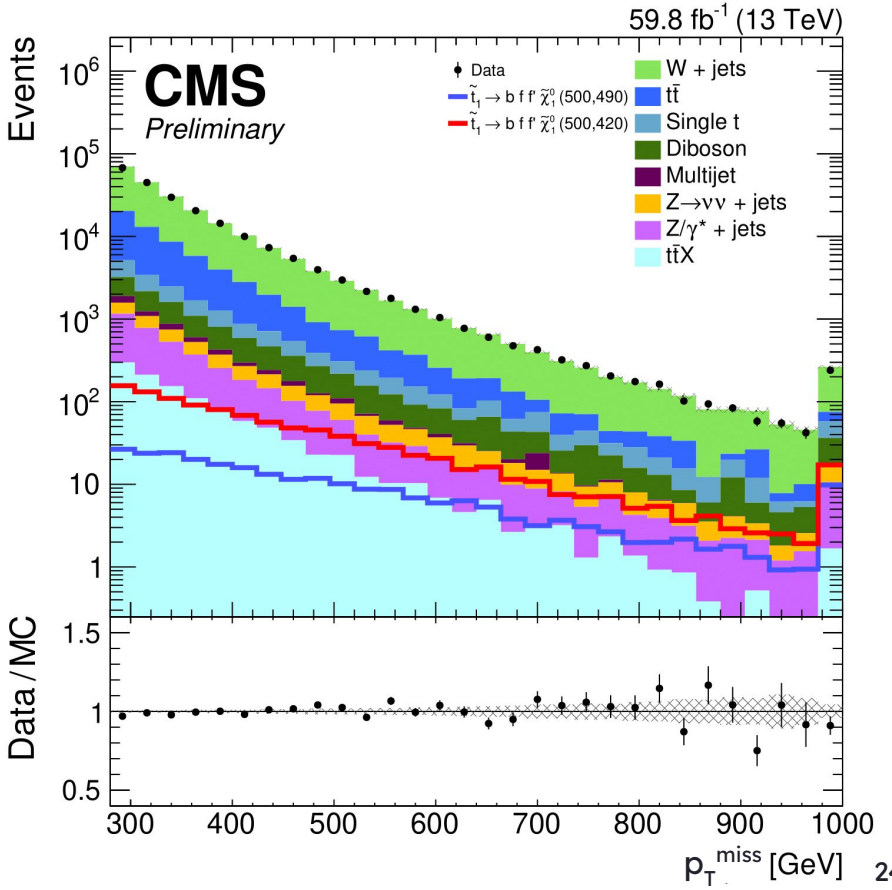
We take advantage of these 2 properties to aggregate the SPs within the same Δm to gain in stats and use them to **train a BDT per Δm** (per year).

Single-lepton training variables

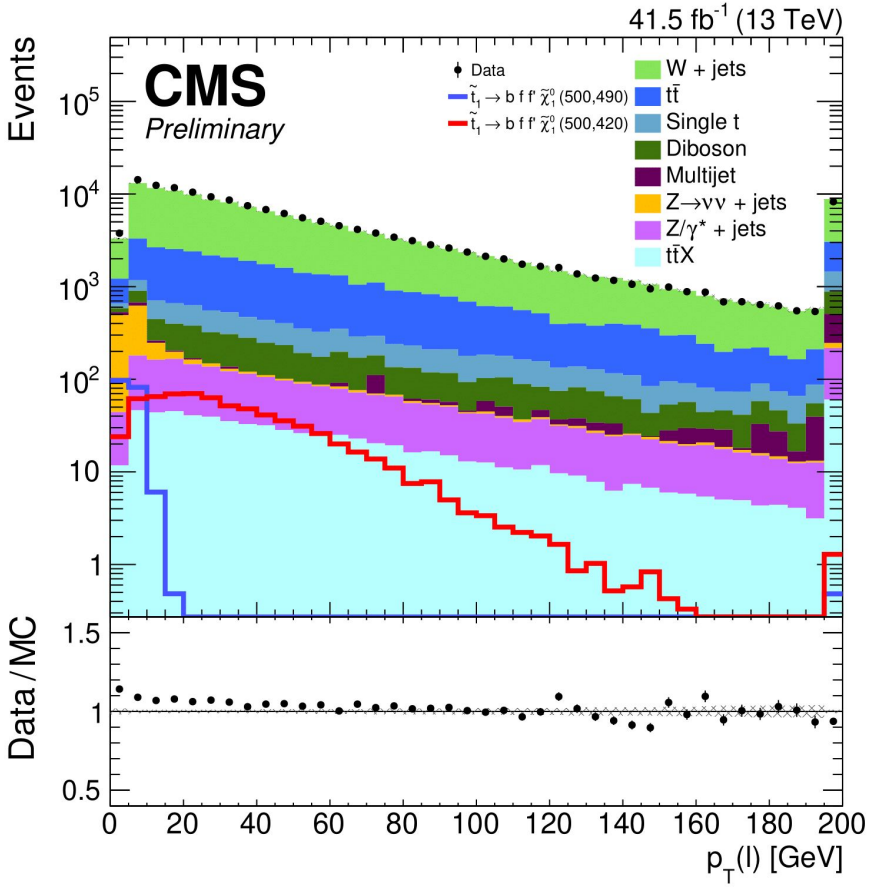
- p_T(lep): lepton (e,μ) p_T
- η(lep): lepton (e,μ) pseudo-rapidity
- chg(lep): lepton (e,μ) charge
- p_T^{miss}: Missing transverse momentum
- M_T: Transverse invariant mass of (p_T^{miss},p_T(lep)) system
- Njets: Number of jets
- $p_T(jet_1)$: Leading jet p_T
- $p_T(jet_b)$: p_T of th jet with the highest b-discriminant
- H_T: scalar sum of the p_T of all jets with $p_T > 30$ GeV
- Njets(b loose): number of loosely b-tagged jets
- ΔR(lep, jet_b): distance between the lepton and the jet with the highest b-discriminant
- Disc(jet_b): b-discriminant distribution of the jet with the highest b-discriminant

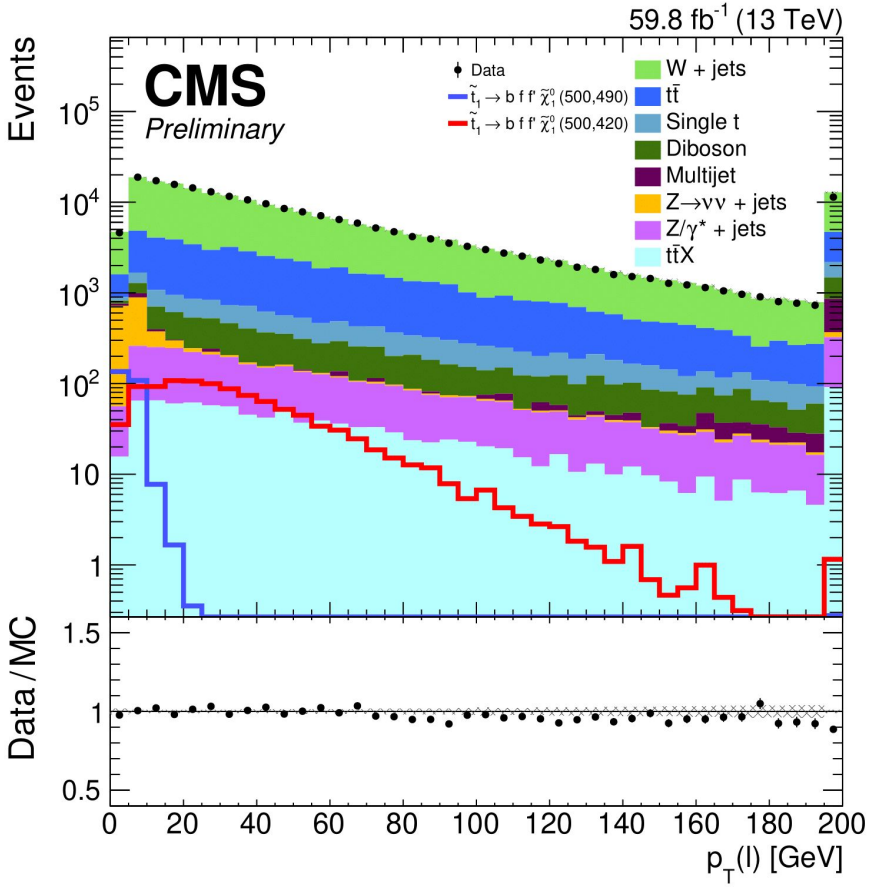
Single-lepton training variables: p_T^{miss}



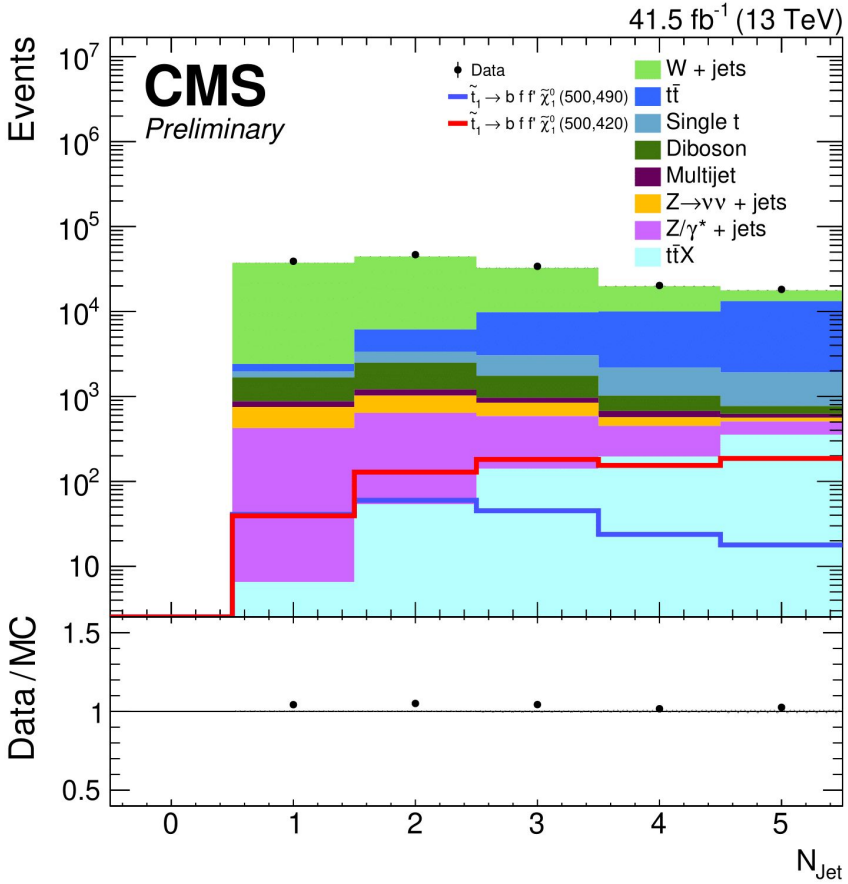


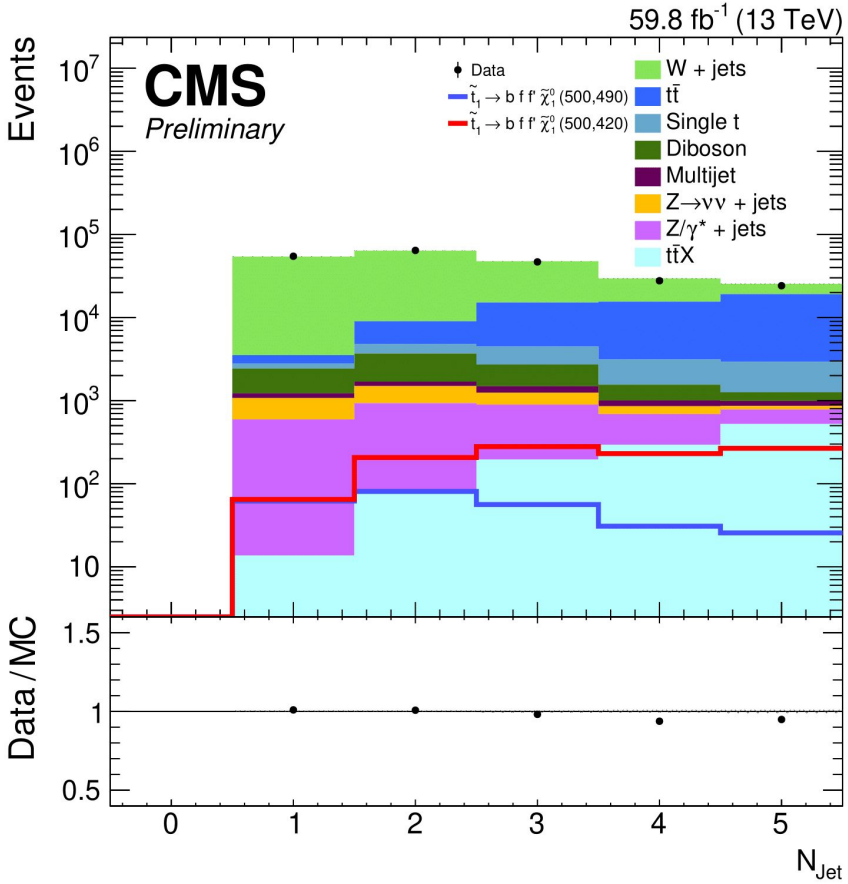
Single-lepton training variables: $p_T(l)$



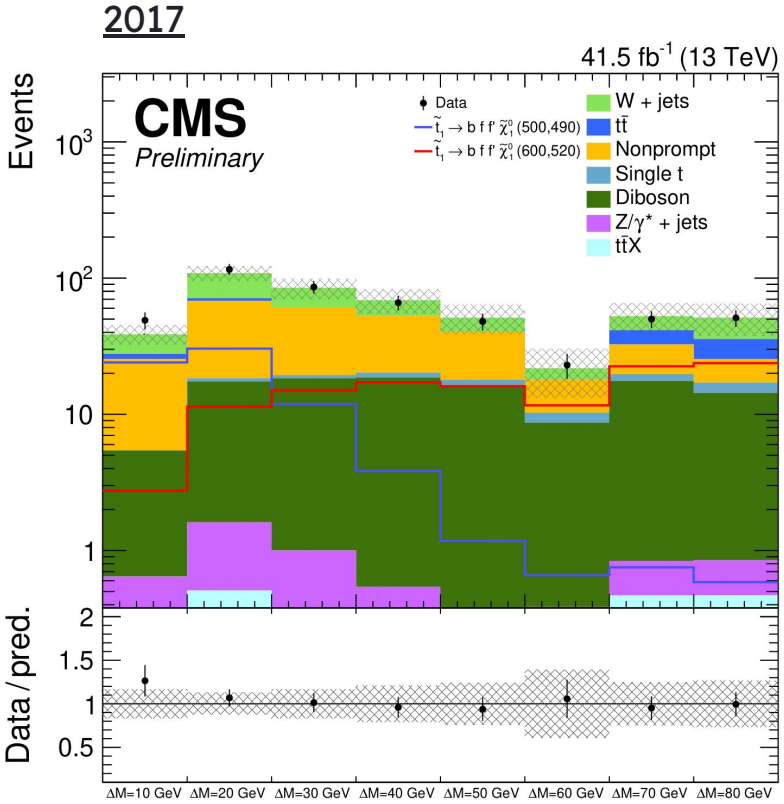


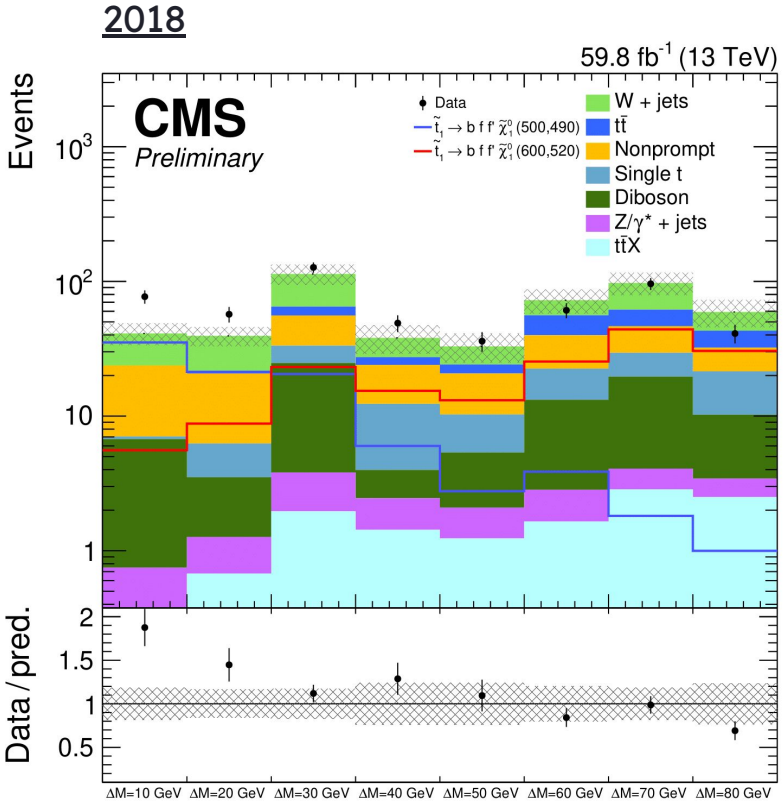
Single-lepton training variables: N_{Jet}





Single-lepton unblinded Signal Region





The 3 main backgrounds (WJets, TTbar and Nonprompt leptons) are estimated from data driven methods