

[MadAnalysis5 Project]
Recasting: CMS-SUS-16-041

Byeonghak Ko, Jubin Park
Kyeongpil Lee and Tae Gyu Ro
with **Guillaume Chalons**

MadAnalysis 5 Workshop
on LHC recasting @ Korea

※ Wiki: <https://madanalysis.irmp.ucl.ac.be/wiki/Korea04-CMS-SUS-16-041>

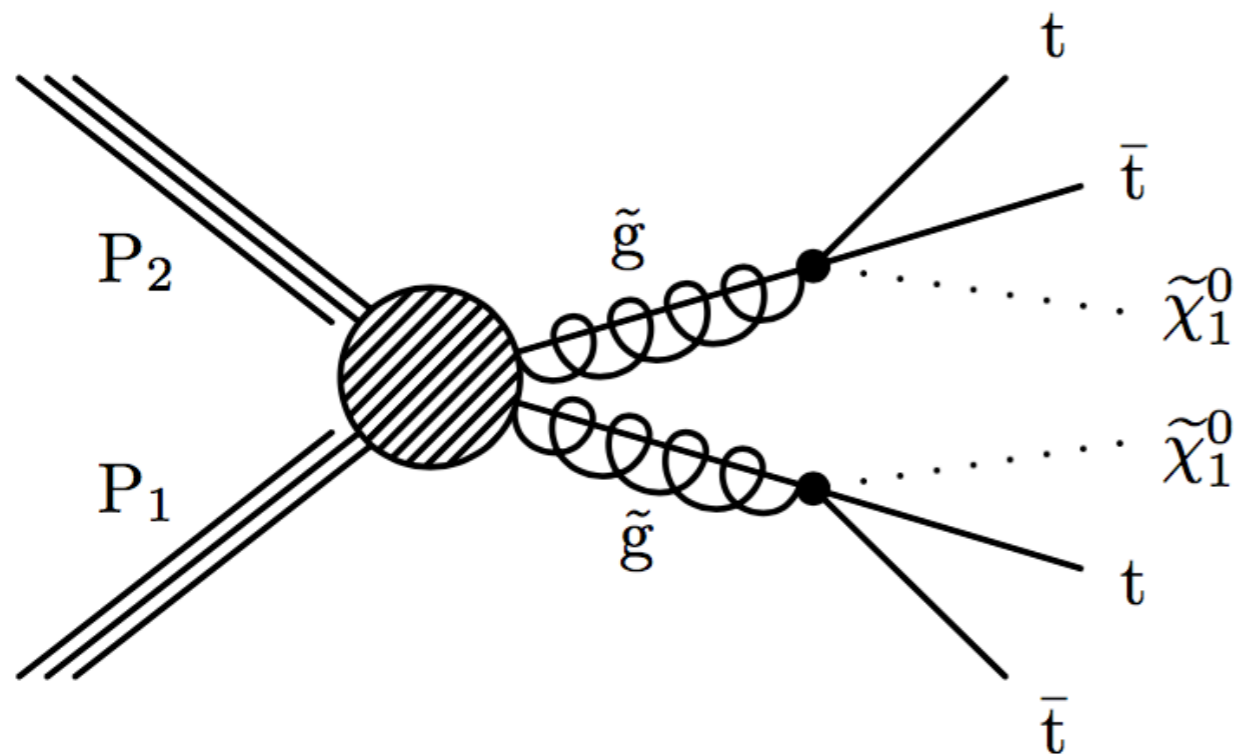
Introduction

- ◉ CMS-SUS-16-041: Search for new physics in events with **multileptons** and **jets** in **35.9 fb⁻¹** of proton-proton collision data at $\sqrt{s} = 13$ TeV
 - New physics search in the final state with **multi-leptons** and **jets**
 - Reference model: Simplified models for LHC^[1]
 - Free parameters: Mass of gluino, squark, and LSP (Lightest SUSY particle)
 - ✓ The other particles are assumed to be too heavy to be discovered at the LHC

Introduction

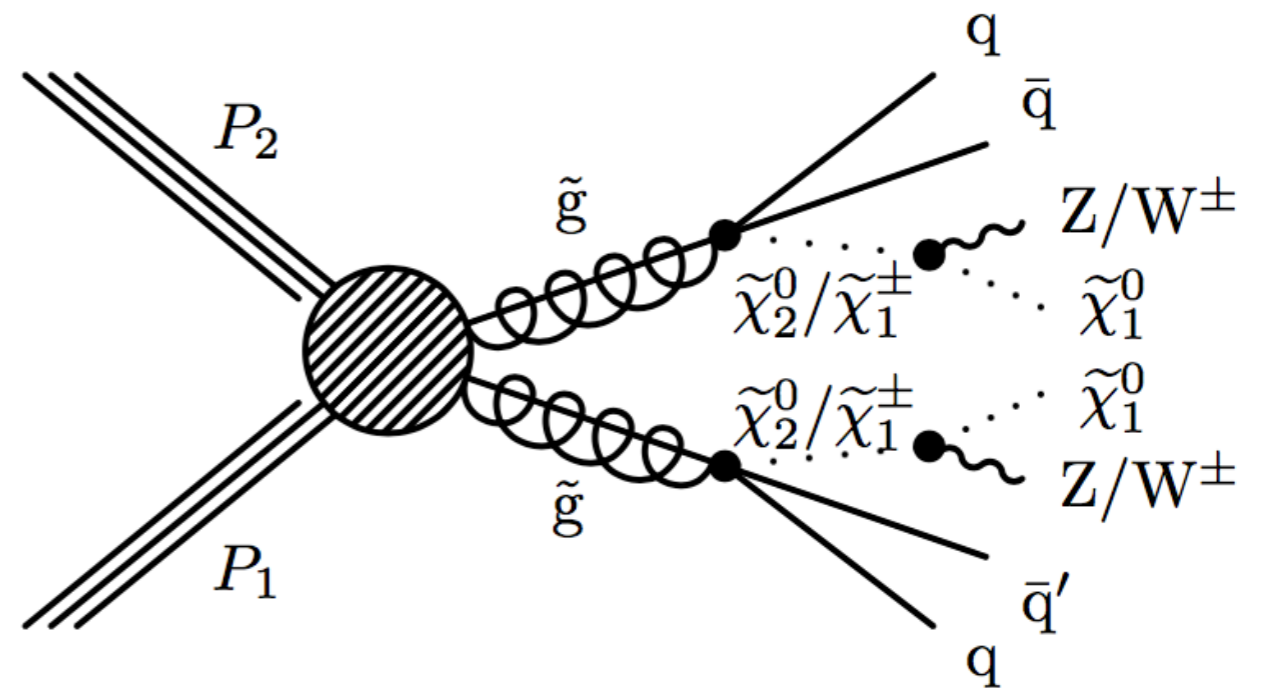
- ◉ CMS-SUS-16-041: Search for new physics in events with **multileptons and jets in 35.9 fb^{-1} of proton-proton collision data at $\sqrt{s} = 13 \text{ TeV}$**
 - New physics search in the final state with **multi-leptons and jets**
 - Reference model: Simplified models for LHC^[1]
 - Free parameters: Mass of gluino, squark, and LSP (Lightest SUSY particle)
 - ✓ The other particles are assumed to be too heavy to be discovered at the LHC

Glauino pair production \rightarrow 4 top-quark & LSP



(a)

Glauino pair production \rightarrow 4 quark & 2 vector bosons

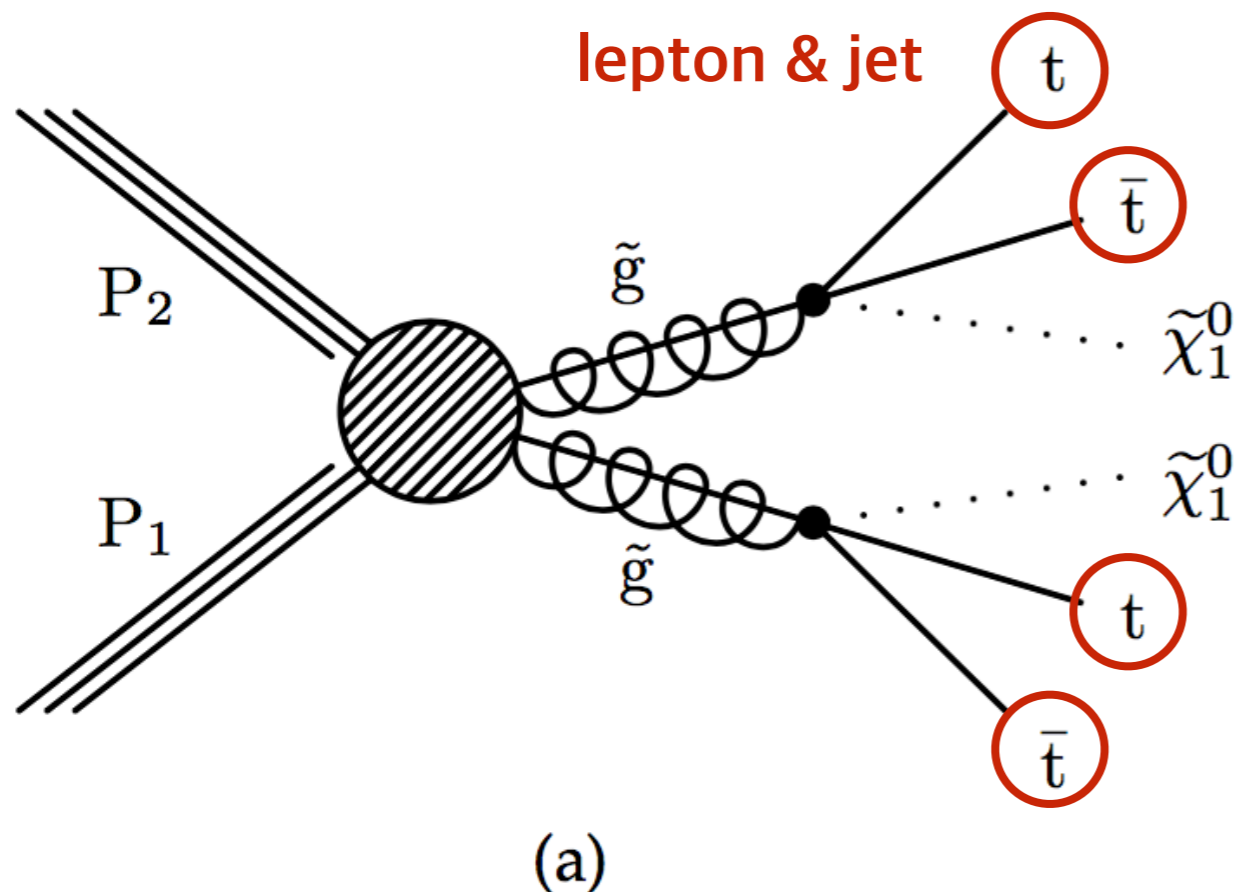


(b)

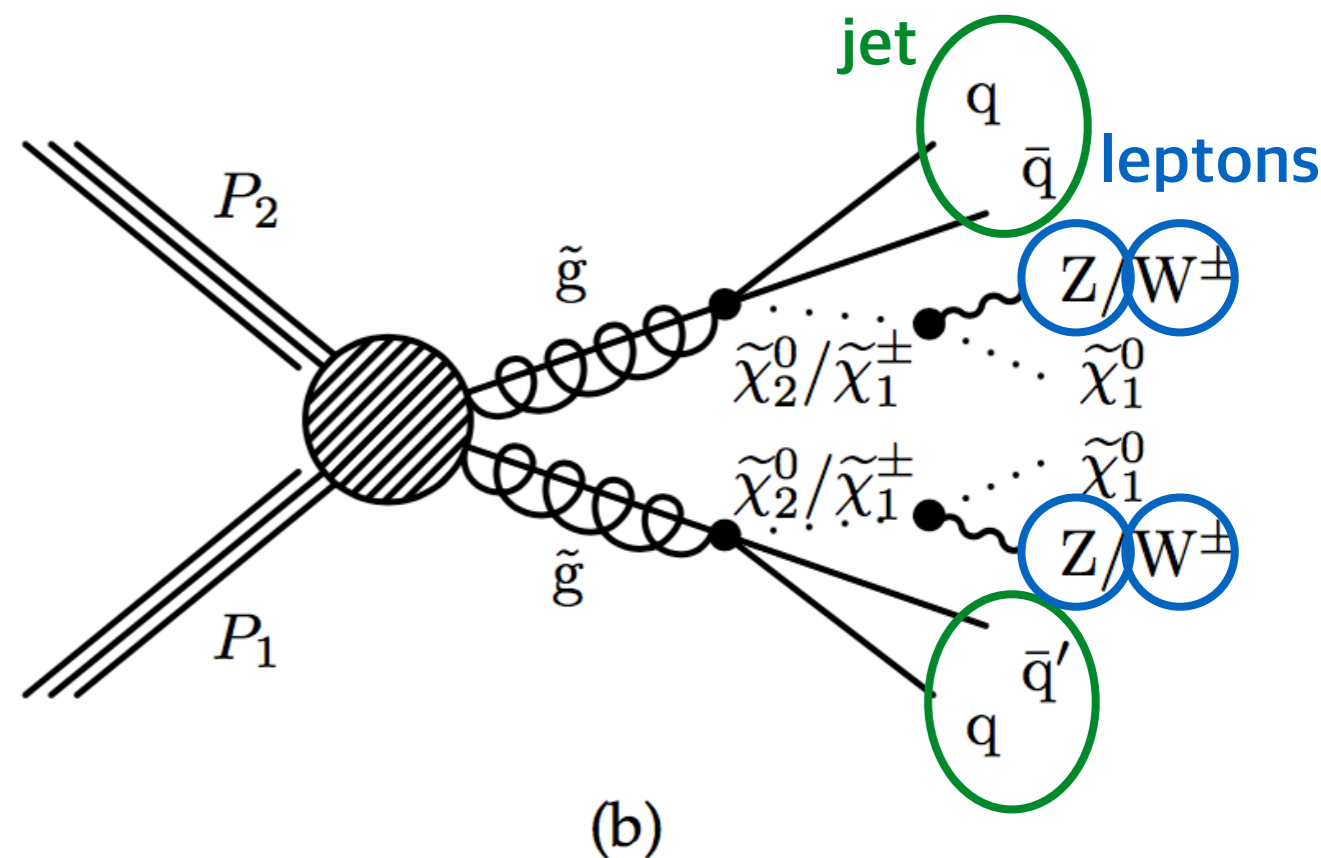
Introduction

- ◉ CMS-SUS-16-041: Search for new physics in events with **multileptons and jets** in **35.9 fb⁻¹** of proton-proton collision data at $\sqrt{s} = 13$ TeV
 - New physics search in the final state with **multi-leptons** and **jets**
 - Reference model: Simplified models for LHC^[1]
 - Free parameters: Mass of gluino, squark, and LSP (Lightest SUSY particle)
 - ✓ The other particles are assumed to be too heavy to be discovered at the LHC

Glauino pair production → 4 top-quark & LSP



Glauino pair production → 4 quark & 2 vector bosons



Introduction

- CMS-SUS-16-041: Search for new physics in events with **multileptons and jets** in 35.9 fb^{-1} of proton-proton collision data at $\sqrt{s} = 13 \text{ TeV}$
 - New physics search in the final state with **multi-leptons** and **jets**
 - Reference model: Simplified models for LHC^[1]
 - Free parameters: Mass of gluino, squark, and LSP (Lightest SUSY particle)
 - ✓ The other particles are assumed to be too heavy to be discovered at the LHC

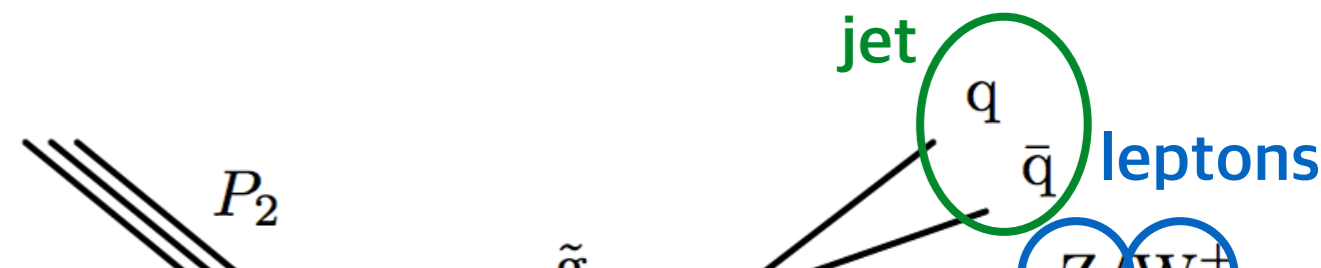
Glauino pair production \rightarrow 4 top-quark & LSP

Glauino pair production \rightarrow 4 quark & 2 vector bosons

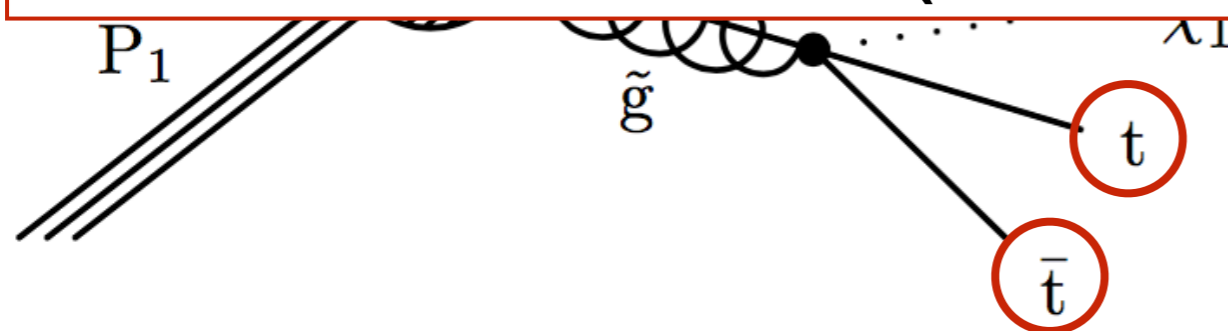
lepton & jet



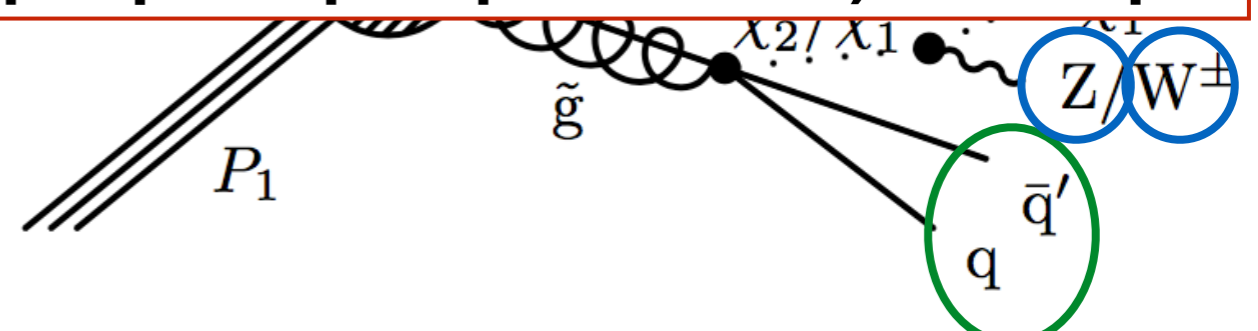
jet



**We studies these two scenarios: Cut-flows are provided
Other 2 more scenarios (bottom & top squark pair production): backup**



(a)



(b)

Introduction

Search strategy

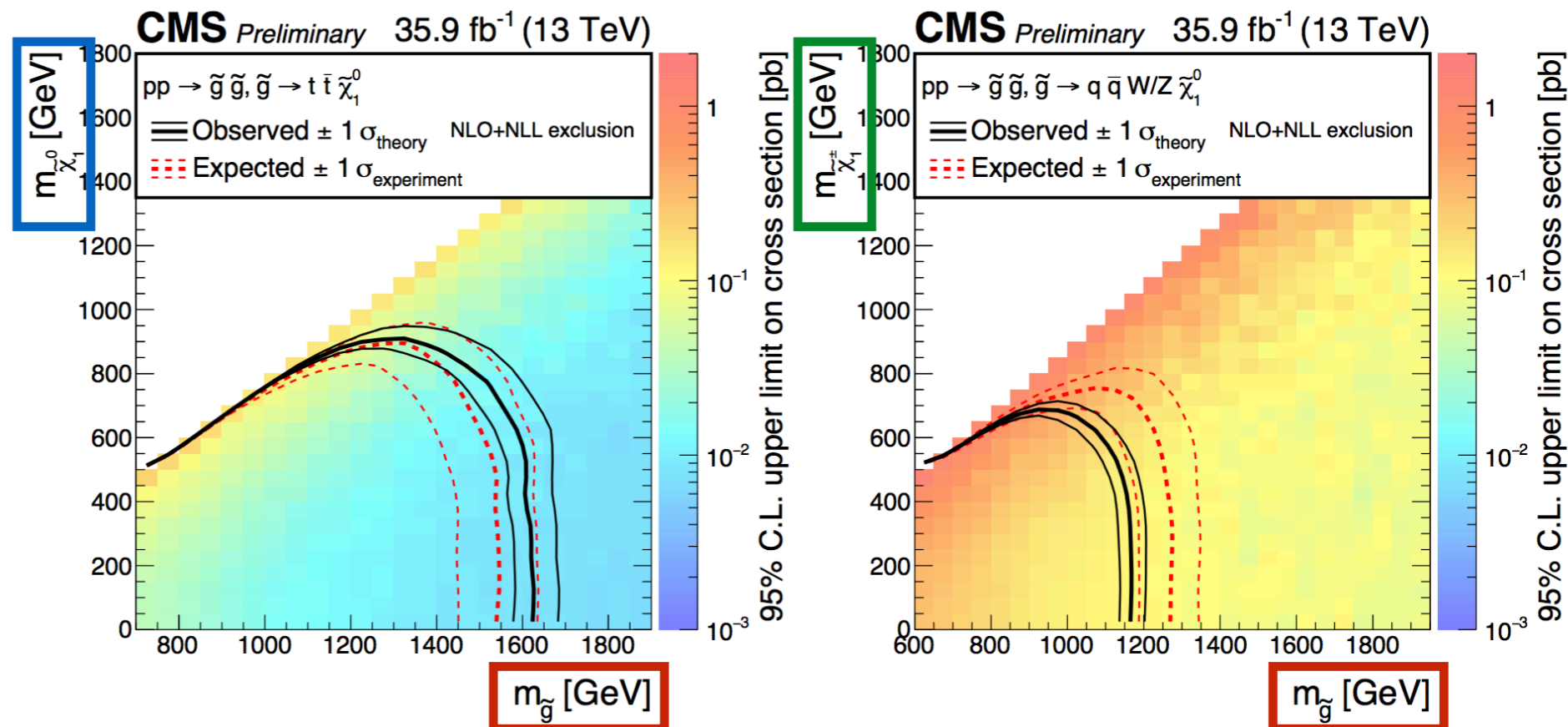
- Divide into several signal region & count the number of events

- N_{jets}
- $N_{\text{b-jets}}$
- H_{T} : Sum of jet P_{T}
- $E_{\text{T}}^{\text{miss}}$
- M_{T} : transverse mass

N_{jets}	$N_{\text{b-jets}}$	H_{T} (GeV)	$50(70) \text{ GeV} \leq E_{\text{T}}^{\text{miss}} < 150 \text{ GeV}$	$150 \text{ GeV} \leq E_{\text{T}}^{\text{miss}} < 300 \text{ GeV}$	$E_{\text{T}}^{\text{miss}} \geq 300 \text{ GeV}$
≥ 2	0	60 – 400	SR1 †	SR2 †	SR16 †
		400 – 600	SR3 †	SR4 †	
	1	60 – 400	SR5	SR6	
		400 – 600	SR7	SR8	
	2	60 – 400	SR9	SR10	
		400 – 600	SR11	SR12	
≥ 3	60 – 600	SR13			
inclusive	≥ 600	SR14 †	SR15 †		

Results

- Excluded region: $m(\text{gluino})$ - $m(\text{LSP})$ plane (and the other variables)



Recasting: Sample Generation

- ◉ Generator level (Madgraph5 + Pythia8)
 - Generate signal samples used for cut-flow check
 - [1] Gluino pair production → **4 top-quark & LSP**
 - ✓ $M(\text{gluino}) = 1500 \text{ GeV}$, $M(\text{LSP}) = 200 \text{ GeV}$ ($\sigma = 0.0142 \text{ pb}$)
 - [2] Gluino pair production → **4 quark & 2 vector bosons**
 - ✓ $M(\text{gluino}) = 1200 \text{ GeV}$, $M(\text{LSP}) = 400 \text{ GeV}$ ($\sigma = 0.0856 \text{ pb}$)
 - Decay of SUSY particle (from effective vertex): in pythia8
 - Generate **background** processes: ttbar, ttZ, ttX, VVV

Recasting: Sample Generation

Generator level (Madgraph5 + Pythia8)

- Generate signal samples used for cut-flow check
 - [1] Gluino pair production → **4 top-quark & LSP**
 - ✓ $M(\text{gluino}) = 1500 \text{ GeV}$, $M(\text{LSP}) = 200 \text{ GeV}$ ($\sigma = 0.0142 \text{ pb}$)
 - [2] Gluino pair production → **4 quark & 2 vector bosons**
 - ✓ $M(\text{gluino}) = 1200 \text{ GeV}$, $M(\text{LSP}) = 400 \text{ GeV}$ ($\sigma = 0.0856 \text{ pb}$)
 - Decay of SUSY particle (from effective vertex): in pythia8
- Generate **background** processes: ttbar, ttZ, ttX, VVV

Reconstruction level (Delphes)

- Default Delphes card used in CMS fast simulation is provided
- **Remove the default isolation conditions on leptons**
 - This affect on cut-flow results
 - Isolation will be applied in the analysis level

Recasting: MadAnalysis 5

- ◉ Implement variable cuts following analysis procedure
 - Baseline selection
 - $N_{\text{jets}} \geq 2$ with $P_T^{\text{jet}} > 30$ GeV, $|\eta| < 2.4$
 - $E_T^{\text{miss}} > 50$ GeV
 - $N_{\text{leptons}} \geq 3$ with
 - ✓ [1] **kinematic cuts** with $P_T > 25, 15, 10$ GeV (depending on H_T & lepton flavor) & $|\eta| < 2.4$

Recasting: MadAnalysis 5

- Implement variable cuts following analysis procedure

- Baseline selection

- $N_{\text{jets}} \geq 2$ with $P_{\text{T}}^{\text{jet}} > 30$ GeV, $|\eta| < 2.4$

- $E_{\text{T}}^{\text{miss}} > 50$ GeV

- $N_{\text{leptons}} \geq 3$ with

- ✓ [1] **kinematic cuts** with $P_{\text{T}} > 25, 15, 10$ GeV (depending on H_{T} & lepton flavor) & $|\eta| < 2.4$

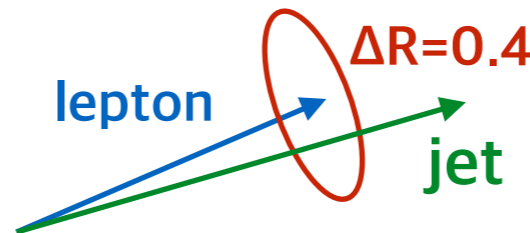
- ✓ [2] **Mini-isolation** $I_{\text{mini}} < 0.16$ (0.12: electron)

- P_{T} -dependent ΔR cone

$$R = \frac{10 \text{ GeV}}{\min(\max(p_{\text{T}}(\ell), 50), 200)}, \Delta R \equiv \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2},$$

- ✓ [3] $P_{\text{T}}^{\text{ratio}} = P_{\text{T}}(\ell)/P_{\text{T}}(\text{jet}) > 0.69$ (0.76)

- Jet: matched to a lepton within $\Delta R = 0.4$



- ✓ [4] $P_{\text{T}}^{\text{rel}} = \text{magnitude of lepton momentum component perpendicular to jet axis} > 6.0$ (7.2)

- ✓ [5] Combination of 3 variables: I_{mini} and $(P_{\text{T}}^{\text{ratio}}$ or $P_{\text{T}}^{\text{rel}})$

$$p_{\text{T}}^{\text{rel}} = \frac{(\vec{p}(\text{jet}) - \vec{p}(\ell)) \cdot \vec{p}(\ell)}{\|\vec{p}(\text{jet}) - \vec{p}(\ell)\|}.$$

Recasting: MadAnalysis 5

- Implement variable cuts following analysis procedure

- Baseline selection

- $N_{\text{jets}} \geq 2$ with $P_{\text{T}}^{\text{jet}} > 30 \text{ GeV}$, $|\eta| < 2.4$

- $E_{\text{T}}^{\text{miss}} > 50 \text{ GeV}$

- $N_{\text{leptons}} \geq 3$ with

- ✓ [1] **kinematic cuts** with $P_{\text{T}} > 25, 15, 10 \text{ GeV}$ (depending on H_{T} & lepton flavor) & $|\eta| < 2.4$

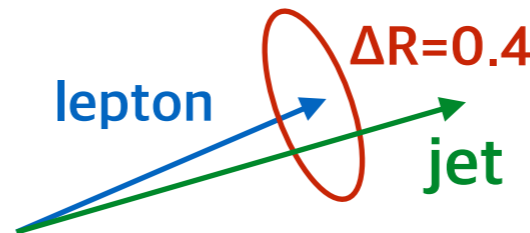
- ✓ [2] **Mini-isolation** $I_{\text{mini}} < 0.16$ (0.12: electron)

- P_{T} -dependent ΔR cone

$$R = \frac{10 \text{ GeV}}{\min(\max(p_{\text{T}}(\ell), 50), 200)}, \Delta R \equiv \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2},$$

- ✓ [3] $P_{\text{T}}^{\text{ratio}} = P_{\text{T}}(\ell)/P_{\text{T}}(\text{jet}) > 0.69$ (0.76)

- Jet: matched to a lepton within $\Delta R = 0.4$



- ✓ [4] $P_{\text{T}}^{\text{rel}} = \text{magnitude of lepton momentum component perpendicular to jet axis} > 6.0$ (7.2)

- ✓ [5] Combination of 3 variables: I_{mini} and ($P_{\text{T}}^{\text{ratio}}$ or $P_{\text{T}}^{\text{rel}}$)

- ✓ [6] **dilepton mass $M(\ell\ell) > 12 \text{ GeV}$** for all possible pairs

$$p_{\text{T}}^{\text{rel}} = \frac{(\vec{p}(\text{jet}) - \vec{p}(\ell)) \cdot \vec{p}(\ell)}{\|\vec{p}(\text{jet}) - \vec{p}(\ell)\|}.$$

Recasting: MadAnalysis 5

Implement variable cuts following analysis procedure

Baseline selection

▸ $N_{\text{jets}} \geq 2$ with $P_{\text{T}}^{\text{jet}} > 30 \text{ GeV}$, $|\eta| < 2.4$

▸ $E_{\text{T}}^{\text{miss}} > 50 \text{ GeV}$

▸ $N_{\text{leptons}} \geq 3$ with

✓ [1] **kinematic cuts** with $P_{\text{T}} > 25, 15, 10 \text{ GeV}$ (depending on H_{T} & lepton flavor) & $|\eta| < 2.4$

✓ [2] **Mini-isolation** $I_{\text{mini}} < 0.16$ (0.12: electron)

- P_{T} -dependent ΔR cone

✓ [3] $P_{\text{T}}^{\text{ratio}} = P_{\text{T}}(\ell)/P_{\text{T}}(\text{jet}) > 0.69$ (0.76)

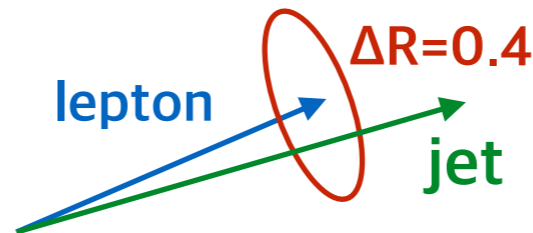
- Jet: matched to a lepton within $\Delta R = 0.4$

✓ [4] $P_{\text{T}}^{\text{rel}} = \text{magnitude of lepton momentum component perpendicular to jet axis} > 6.0$ (7.2)

✓ [5] Combination of 3 variables: I_{mini} and ($P_{\text{T}}^{\text{ratio}}$ or $P_{\text{T}}^{\text{rel}}$)

✓ [6] **dilepton mass** $M(\ell\ell) > 12 \text{ GeV}$ for all possible pairs

$$R = \frac{10 \text{ GeV}}{\min(\max(p_{\text{T}}(\ell), 50), 200)}, \Delta R \equiv \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2},$$



$$p_{\text{T}}^{\text{rel}} = \frac{(\vec{p}(\text{jet}) - \vec{p}(\ell)) \cdot \vec{p}(\ell)}{\|\vec{p}(\text{jet}) - \vec{p}(\ell)\|}.$$

Divide signal region into two categories

• [1] **on-Z**: a lepton pair exists within $76 < M(\ell\ell) < 116$

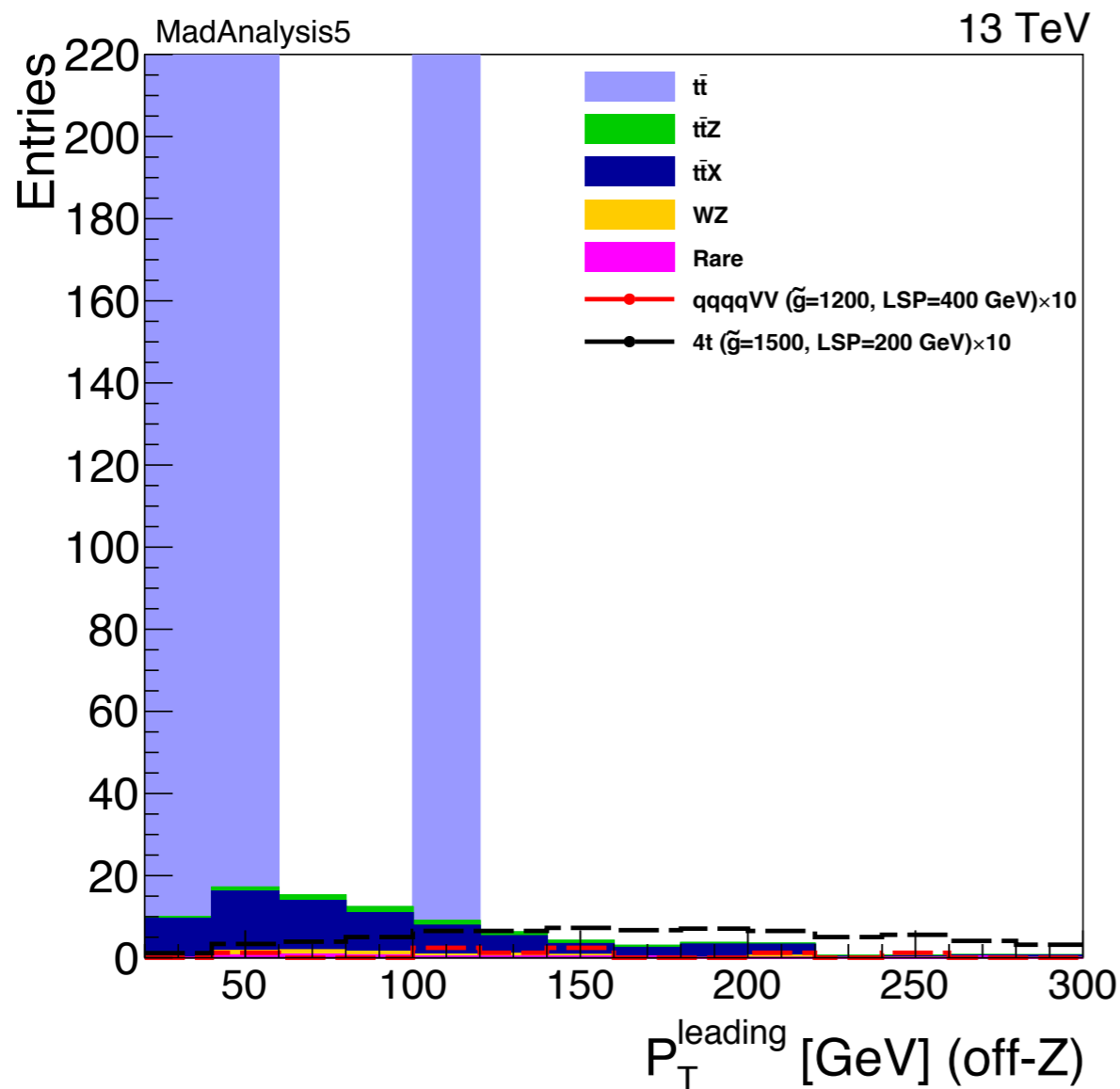
▸ Higher $E_{\text{T}}^{\text{miss}}$ cut to reject Drell-Yan events

• [2] **off-Z**: otherwise

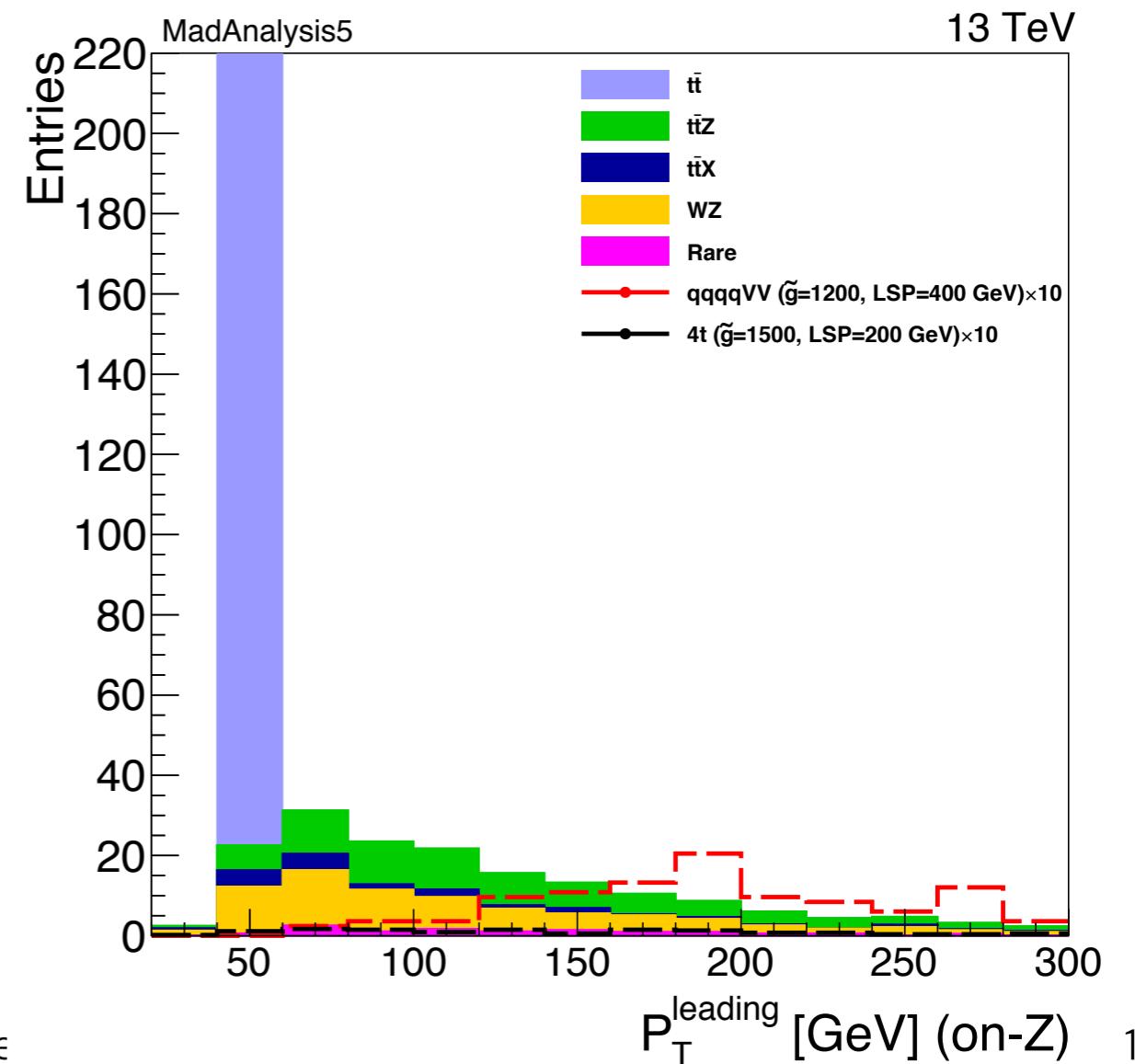
Variable Distribution

- After baseline selection (normalized to 35.9 fb^{-1})
 - Too large contribution from $t\bar{t}$: low statistics & large normalization factor
 - No physical meaning because of statistical fluctuation → remove it
 - This contribution was estimated by data-driven method

Leading lepton P_T (off-Z)

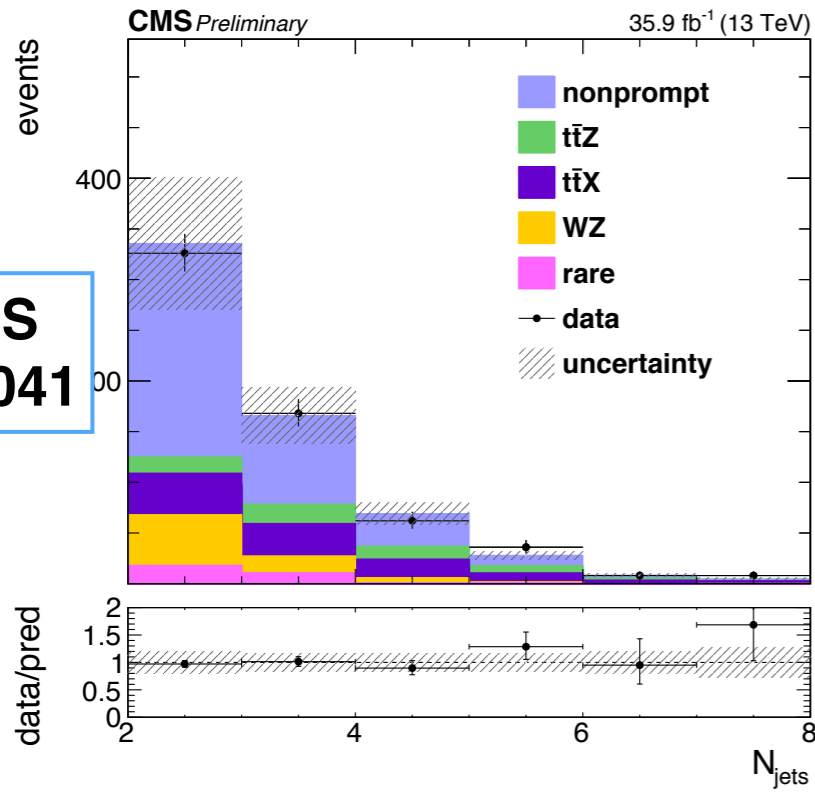


Leading lepton P_T (on-Z)

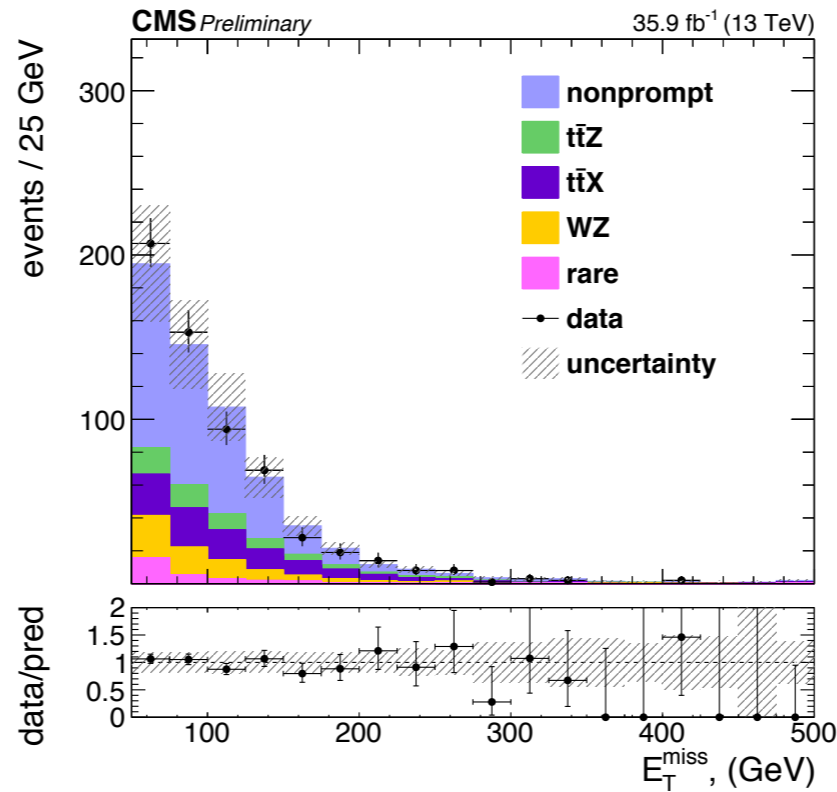


Variable Distribution (off-Z)

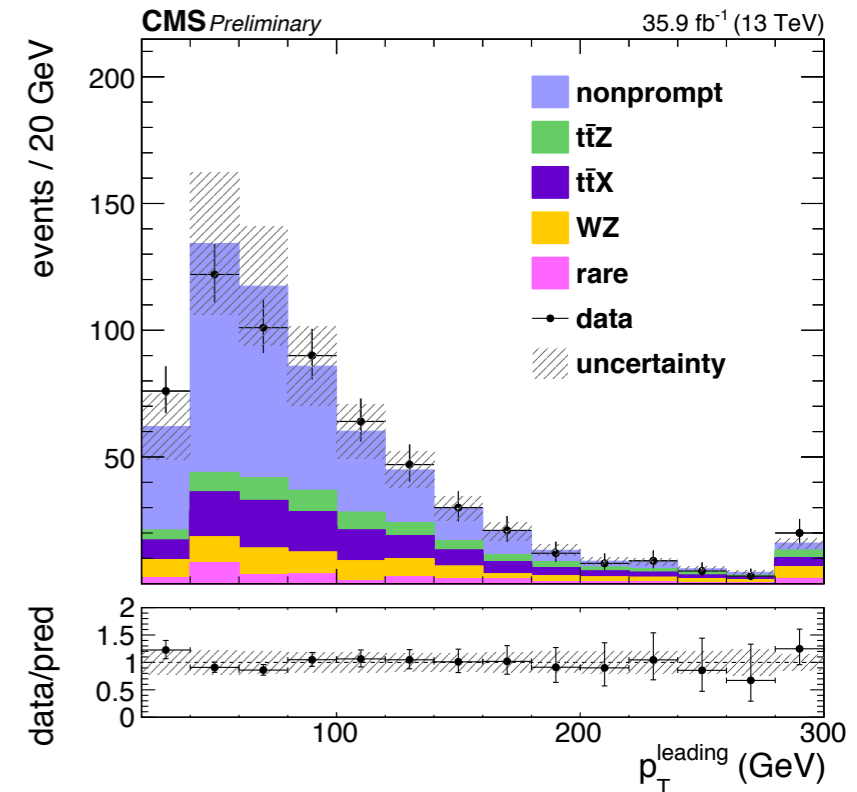
N_{jets}



$E_{\text{T}}^{\text{miss}}$

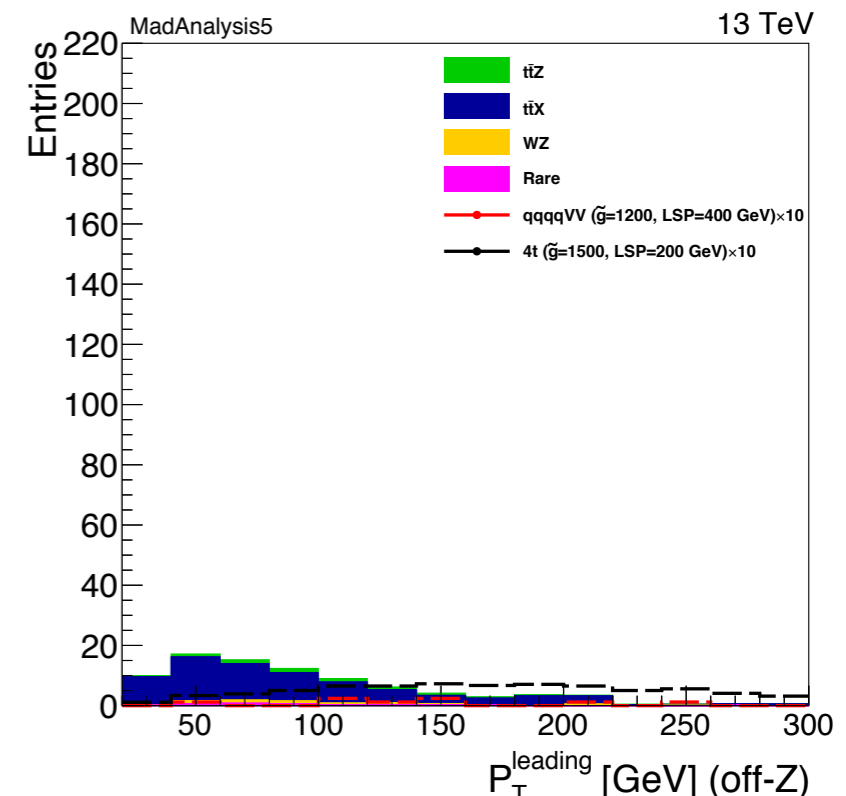
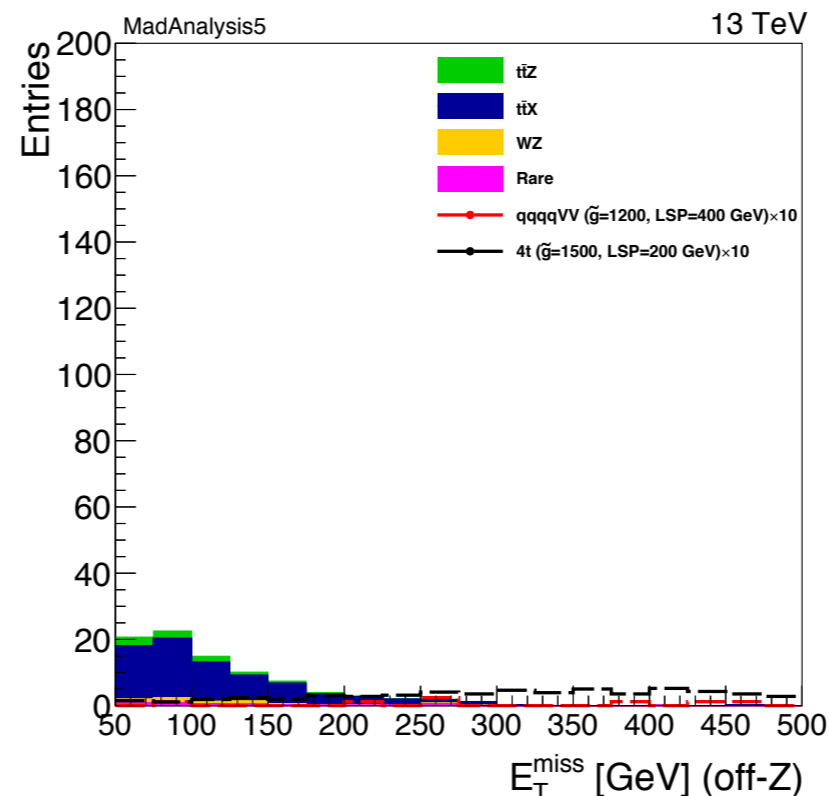
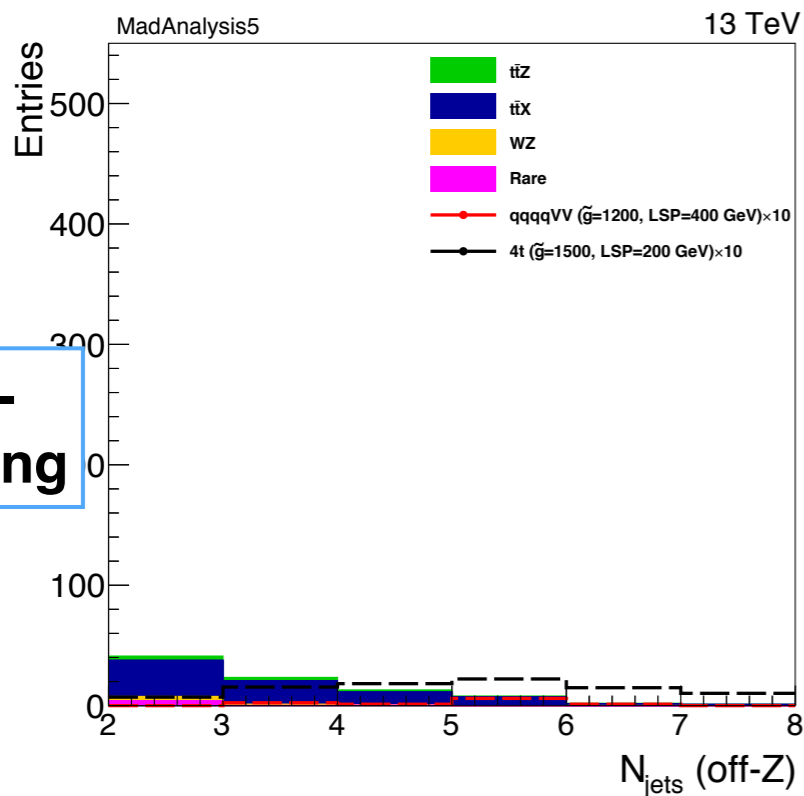


Leading lepton P_{T} (off-Z)



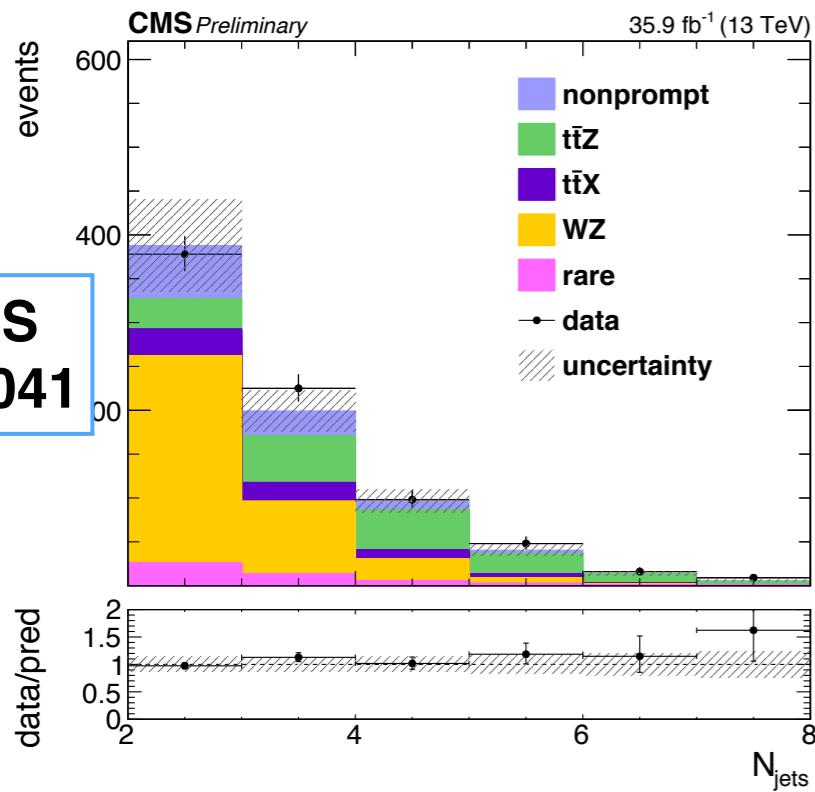
**SUS
-16-041**

**Re-
casting**

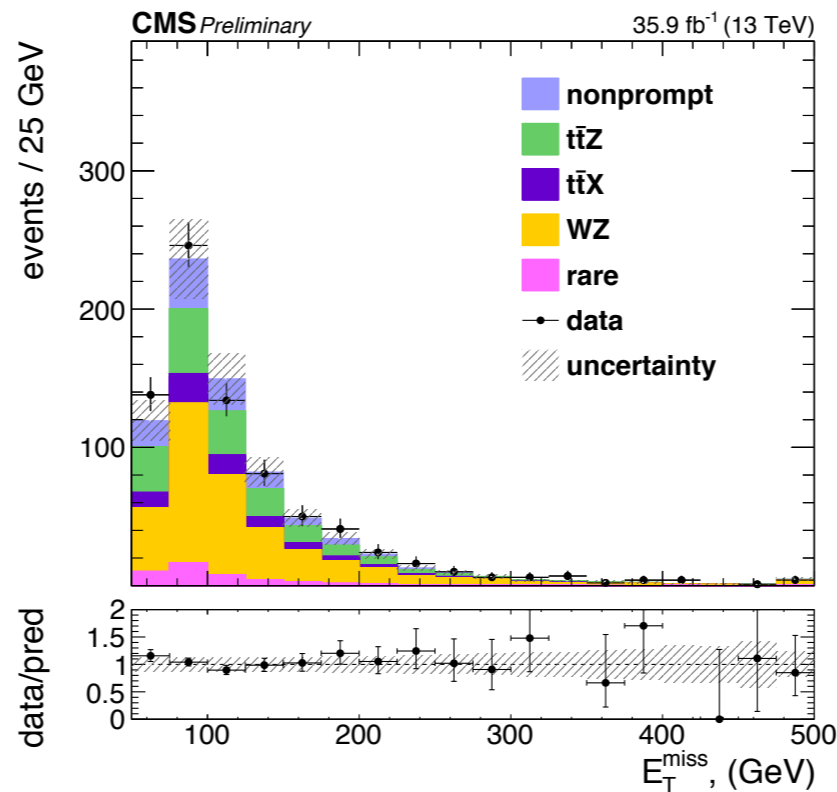


Variable Distribution (on-Z)

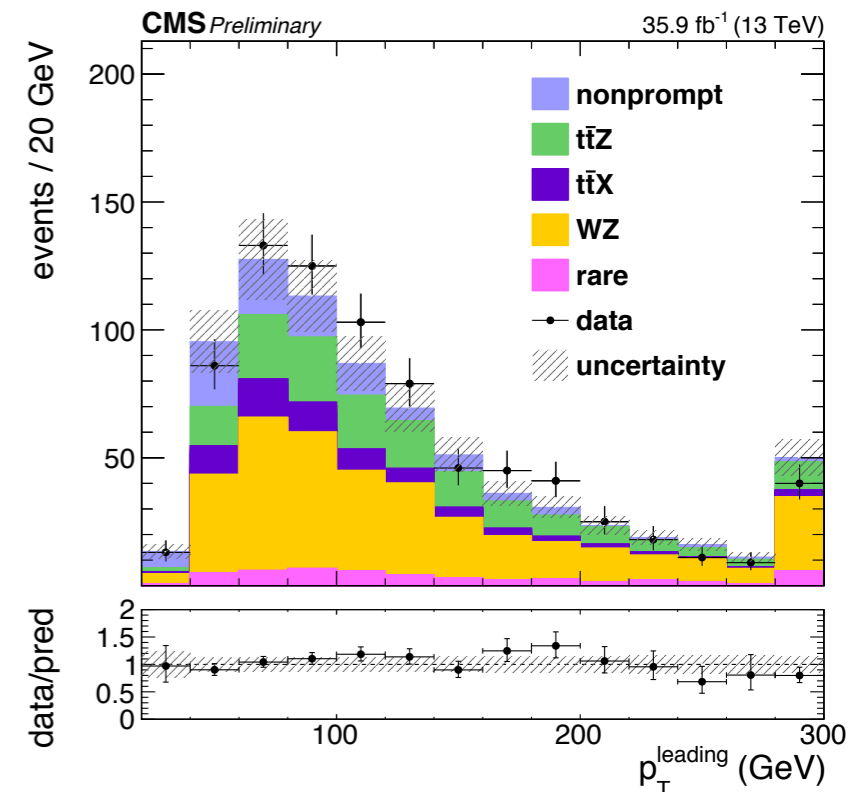
N_{jets}



$E_{\text{T}}^{\text{miss}}$

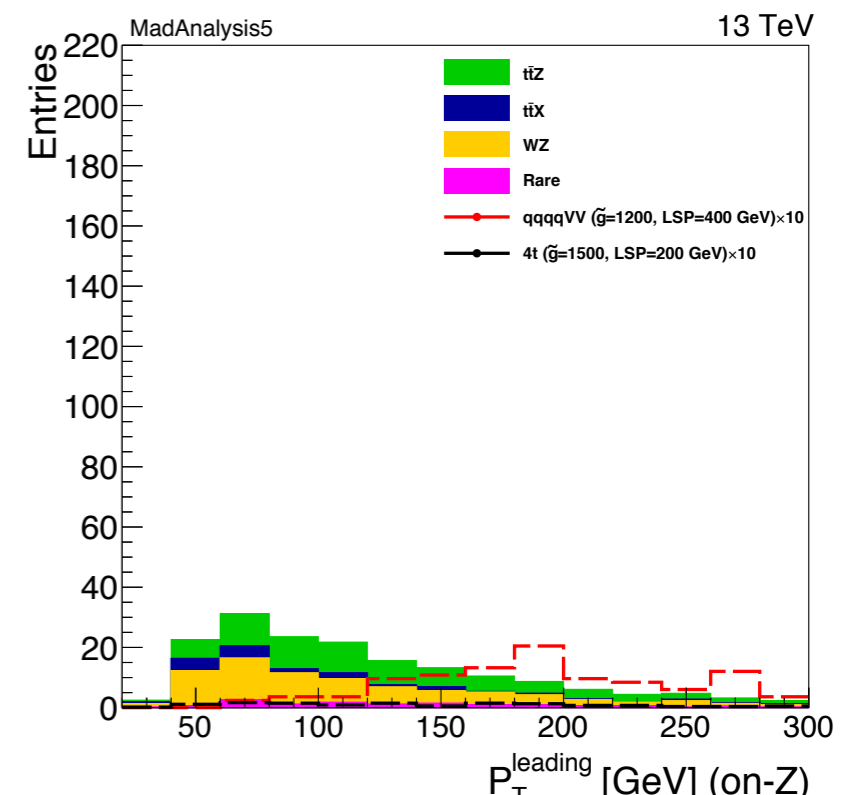
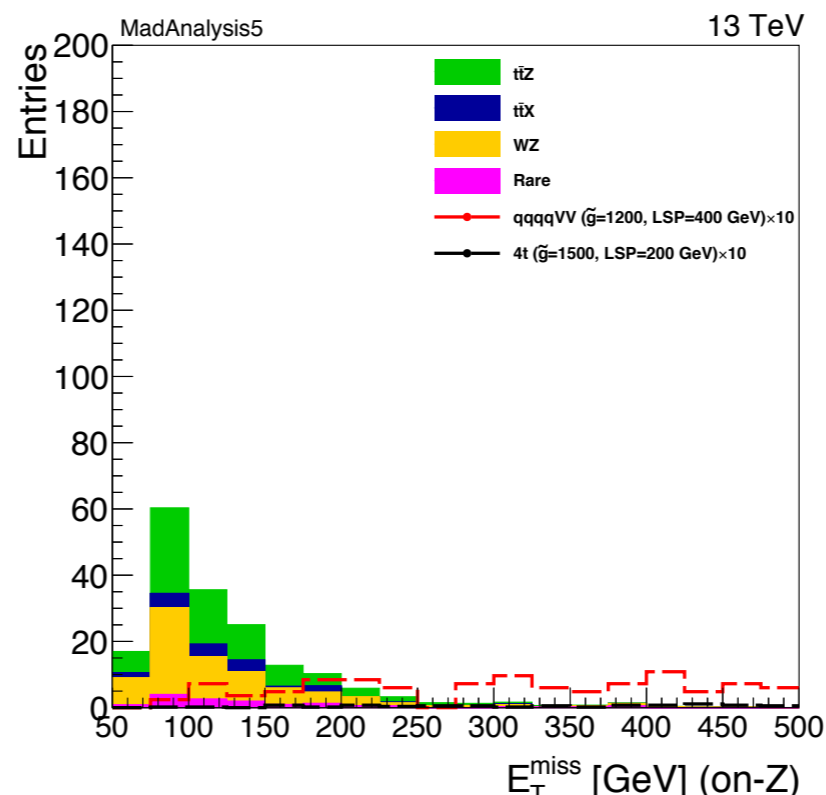
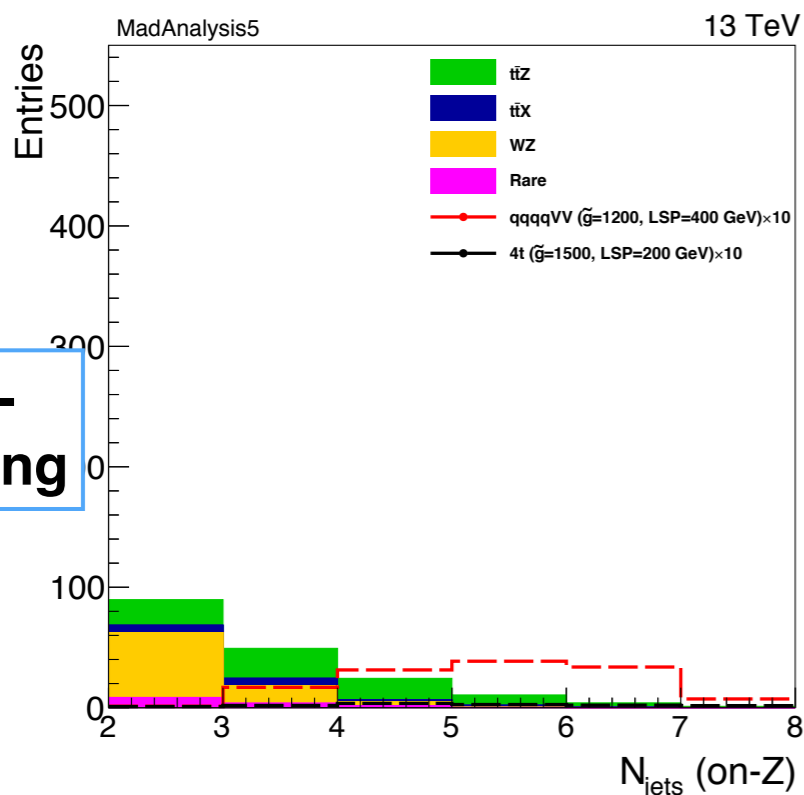


Leading lepton P_{T} (off-Z)



SUS
-16-041

Re-
casting



Cut-flow

- Efficiency = # events passing cuts / # total generated events
- Results with **10k signal events**
 - 4t sample: ~80% disagreement
 - 4qVV sample: < 20% agreement! good!

efficiency: 4t sample

efficiency: 4qVV sample

Cut-flow	Official	Recasting	Difference
Trigger, $N_{\text{lepton}} \geq 3$	1.32%	2.41%	82%↑
$N_{\text{jet}} \geq 2$	1.32%	2.41%	82%↑
$E_{\text{T}}^{\text{miss}} > 50 \text{ GeV}$	1.32%	2.41%	82%↑
Off-Z	1.18%	1.95%	65%↑
SR16a	0.35%	0.99%	283%↑
SR16b	0.49%	0.50%	2%↑

Cut-flow	Official	Recasting	Difference
Trigger, $N_{\text{lepton}} \geq 3$	0.31%	0.35%	13%↑
$N_{\text{jet}} \geq 2$	0.31%	0.35%	13%↑
$E_{\text{T}}^{\text{miss}} > 50 \text{ GeV}$	0.31%	0.35%	13%↑
on-Z	0.30%	0.35%	17%↑
SR15b	0.04%	0.04%	0%
SR16b	0.17%	0.16%	6%↓

Cut-flow

○ Efficiency = # events passing cuts / # total generated events

○ Results with **100k signal events**

- 4t sample: ~60% disagreement
- 4qVV sample: ~60% disagreement :(

→ Is something common factor missing?

efficiency: 4t sample

efficiency: 4qVV sample

Cut-flow	Official	Recasting	Difference
Trigger, $N_{\text{lepton}} \geq 3$	1.32%	2.21%	67%↑
$N_{\text{jet}} \geq 2$	1.32%	2.18%	65%↑
$E_{\text{T}}^{\text{miss}} > 50 \text{ GeV}$	1.32%	2.16%	64%↑
Off-Z	1.18%	1.88%	59%↑
SR16a	0.35%	0.60%	71%↑
SR16b	0.49%	0.78%	59%↑

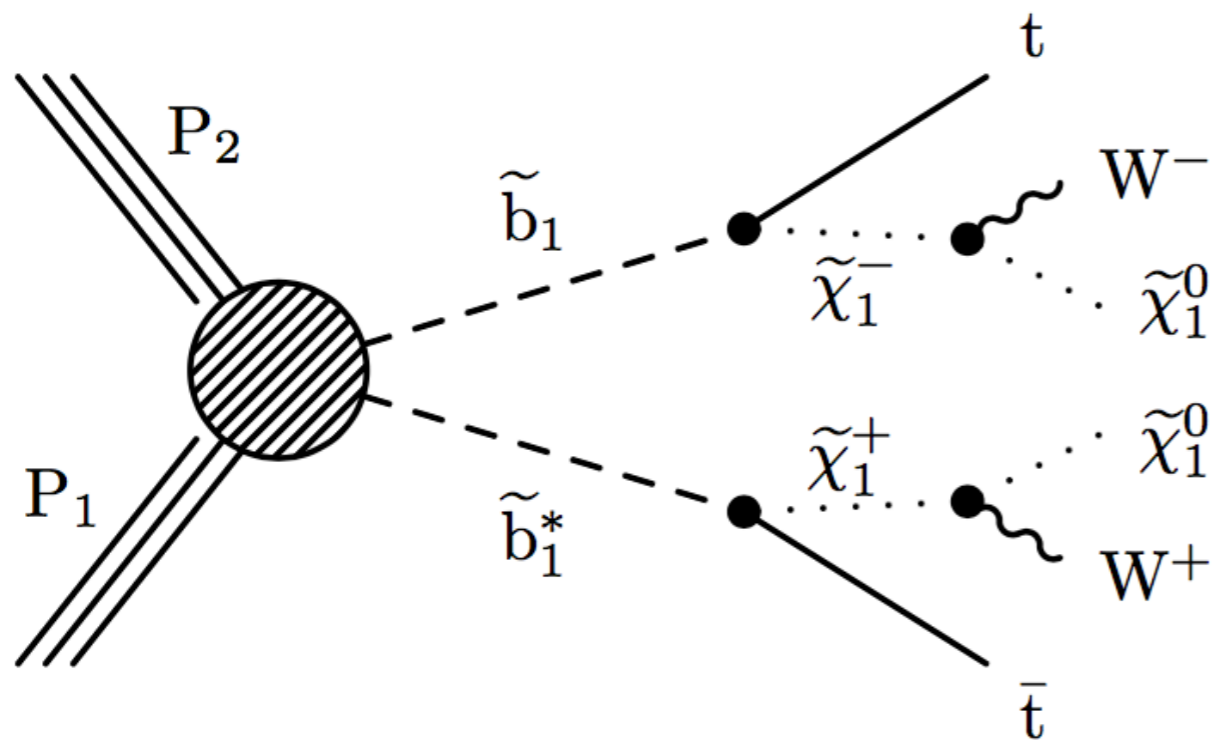
Cut-flow	Official	Recasting	Difference
Trigger, $N_{\text{lepton}} \geq 3$	0.31%	0.52%	68%↑
$N_{\text{jet}} \geq 2$	0.31%	0.52%	68%↑
$E_{\text{T}}^{\text{miss}} > 50 \text{ GeV}$	0.31%	0.50%	61%↑
on-Z	0.30%	0.46%	53%↑
SR15b	0.04%	0.06%	50%↑
SR16b	0.17%	0.26%	53%↑

Summary

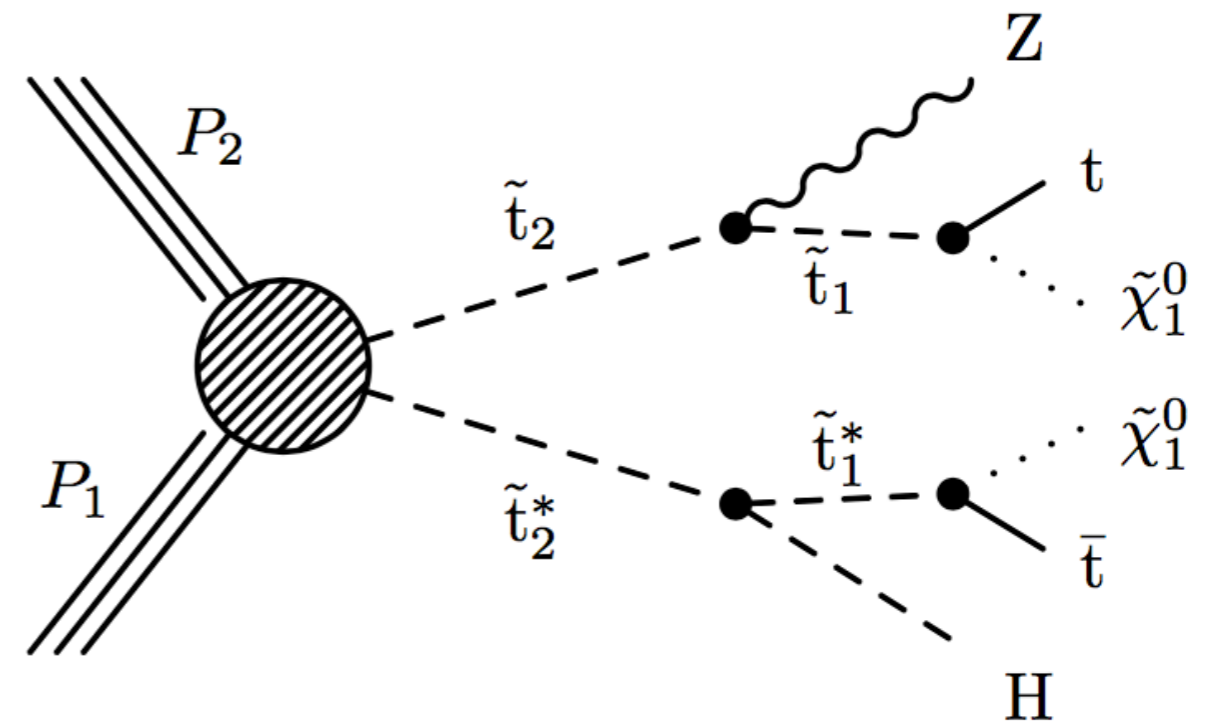
- ◉ Recasting CMS-SUS-16-041: new physics search with **multi-lepton & jets** final state
- ◉ **Sample generation** for signal & background is **successfully done**
- ◉ Variable distribution: Smaller # events, but shape & relative bkg. contribution look similar with the official one
- ◉ Cut-flow: Disagreement by ~60% for both 4t and 4qVV case
 - Still under investigation
 - Follow b-tagging efficiency of CMS paper - but no improvement is observed

Backup

Bottom & Top squark Pair Production



(c)



(d)

Lepton P_T Cuts

```
double PtCut_1st = 0;
double PtCut_2nd = 0;
double PtCut_3rd = 0;
if( HT < 300 )
{
    PtCut_1st = 25;

    // -- sub-leading cut: depend on the lepton flavor -- //
    if( baseLeptons[1]->isMuon() ) PtCut_2nd = 10;
    if( baseLeptons[1]->isElectron() ) PtCut_2nd = 15;

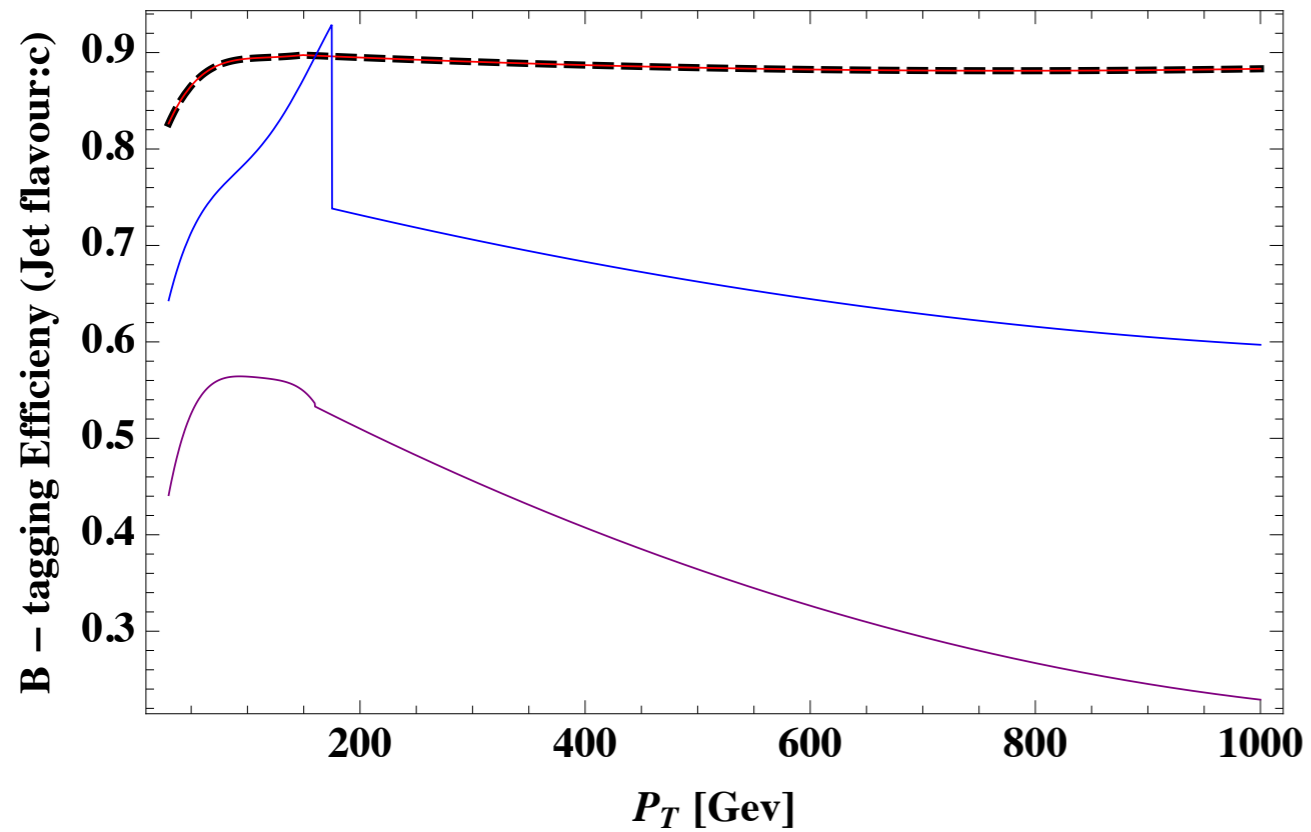
    PtCut_3rd = 10;
}
else if( HT > 300 )
{
    //-- leading cut: depend on the lepton flavor -- //
    if( baseLeptons[0]->isMuon() ) PtCut_1st = 10;
    if( baseLeptons[0]->isElectron() ) PtCut_1st = 15;

    //-- leading cut: depend on the lepton flavor -- //
    if( baseLeptons[1]->isMuon() ) PtCut_2nd = 10;
    if( baseLeptons[1]->isElectron() ) PtCut_2nd = 15;

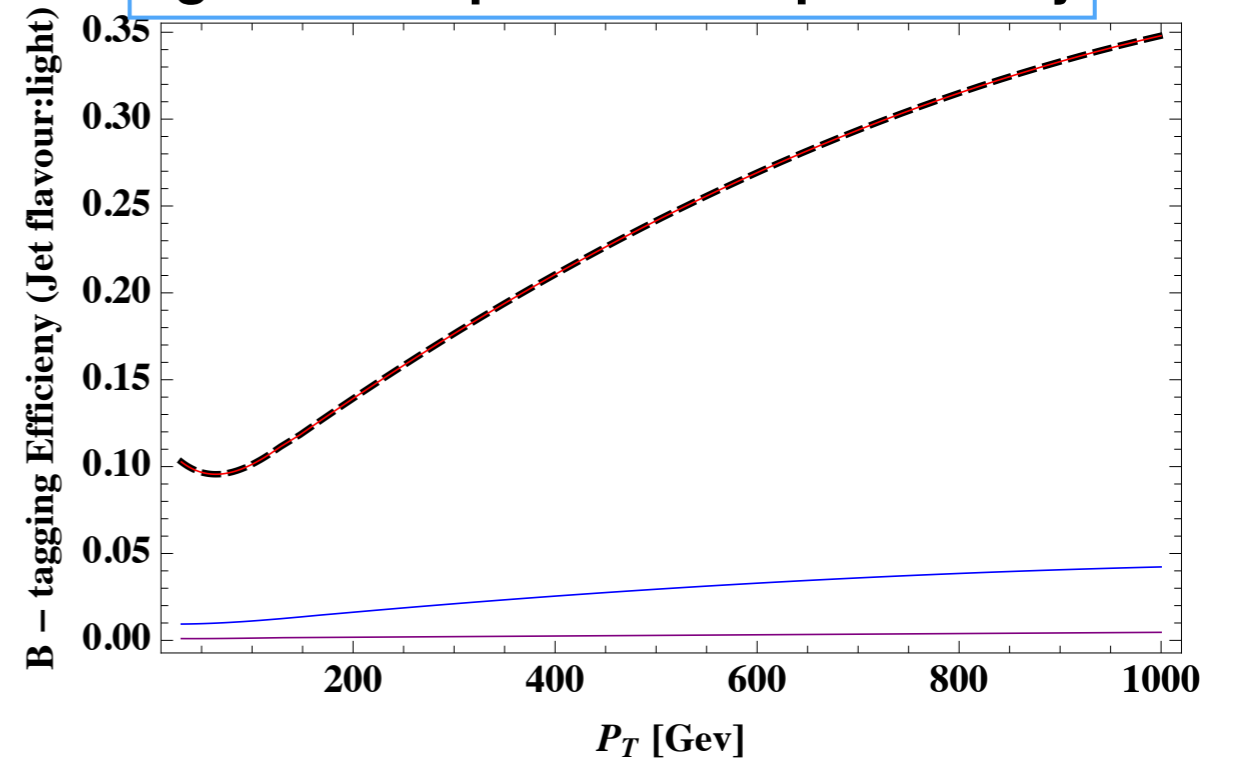
    PtCut_3rd = 10;
}
```

b-tagging Efficiency as a Function of P_T

b-tagging efficiency



light-flavor quark misid. probability



c-quark misid. probability

