Sensitivity study of a search for a simplified supersymmetric model using *pp* collisions at $\sqrt{s} = 13$ TeV with the CMS detector

Tamás Álmos VÁMI

Advisor: Petar MAKSIMOVIC

JOHNS HOPKINS UNIVERSITY

Research Exam

Need for New Physics

We know that the Standard Model (SM) cannot be the final theory.

It does not explain the matter/antimatter asymmetry.

It does not give any candidate on Dark Matter.

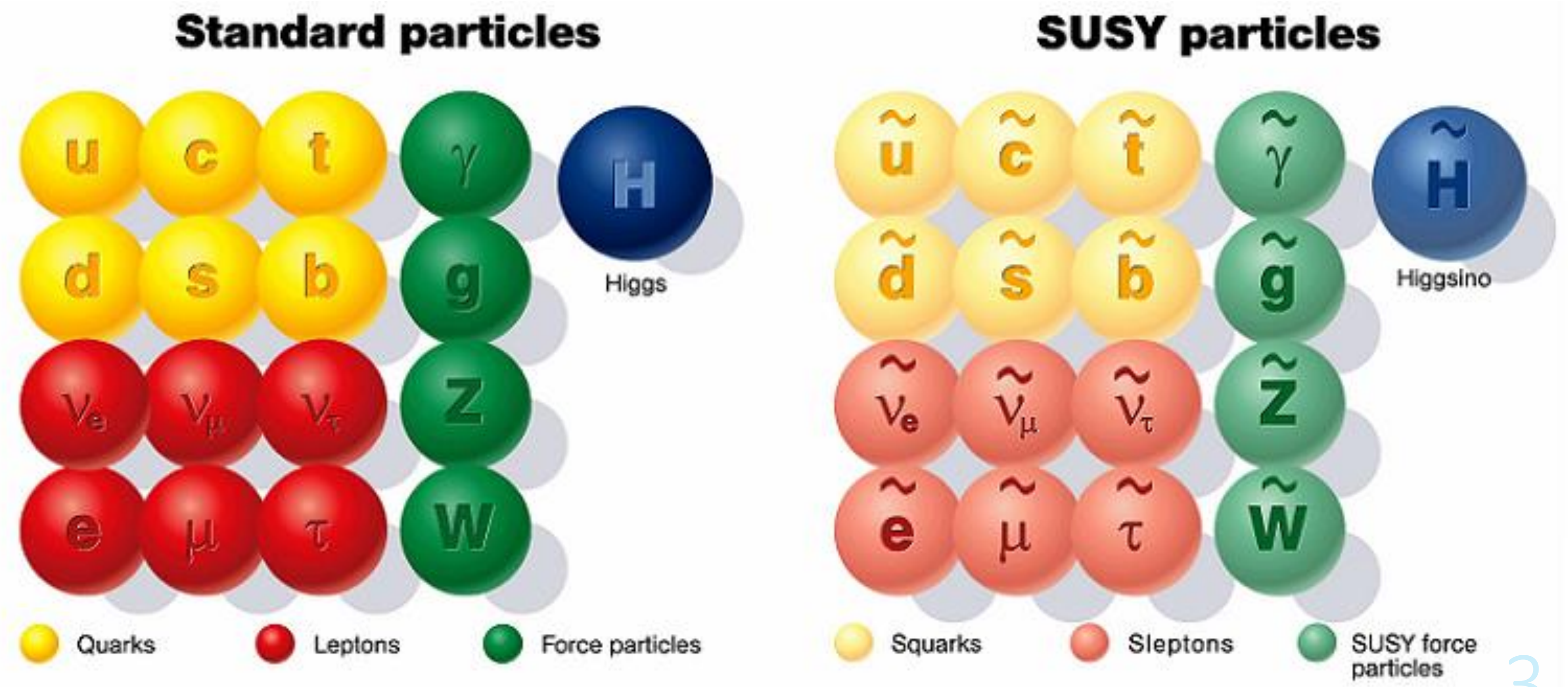
It does not give any reason why the mass of the Higgs is so low wrt the Planck mass.

Unification of the electroweak and the strong interaction is not possible with it.

A possible solution: supersymmetry (SUSY)

Extending the SM with the superpartners of its particles seems to solve these problems theoretically.

Superpartners of the SM fermions are bosons while the SM bosons become fermions.



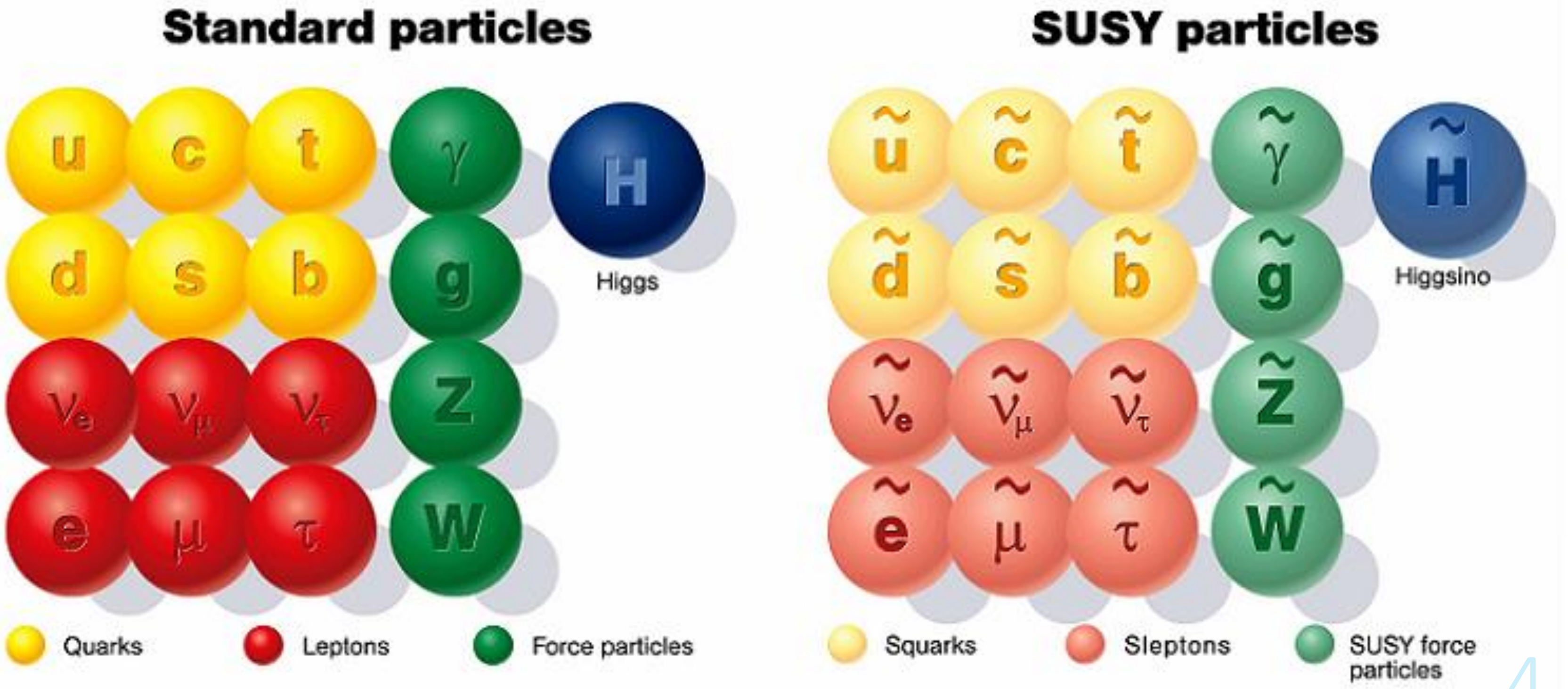
A possible solution: supersymmetry (SUSY)

Extending the SM with the superpartners of its particles seems to solve these problems theoretically.

Superpartners of the SM fermions are bosons while the SM bosons become fermions.

SM gluon → gluino ($\tilde{g}$)
SM quark → squark

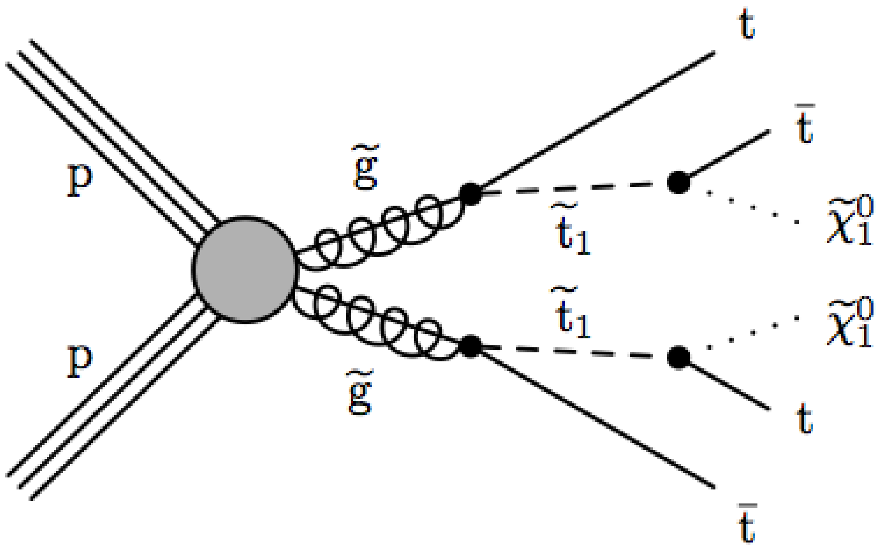
SM bosons → charginos ($\chi^{\pm}_{1,2}$) or neutralinos ($\chi^{0}_{1,2,3,4}$)



Simplified models and boosted objects

Simplified models (SMSs) are effective models with the minimal particle content necessary to produce SUSY-like final states contributing to the channels of interest.

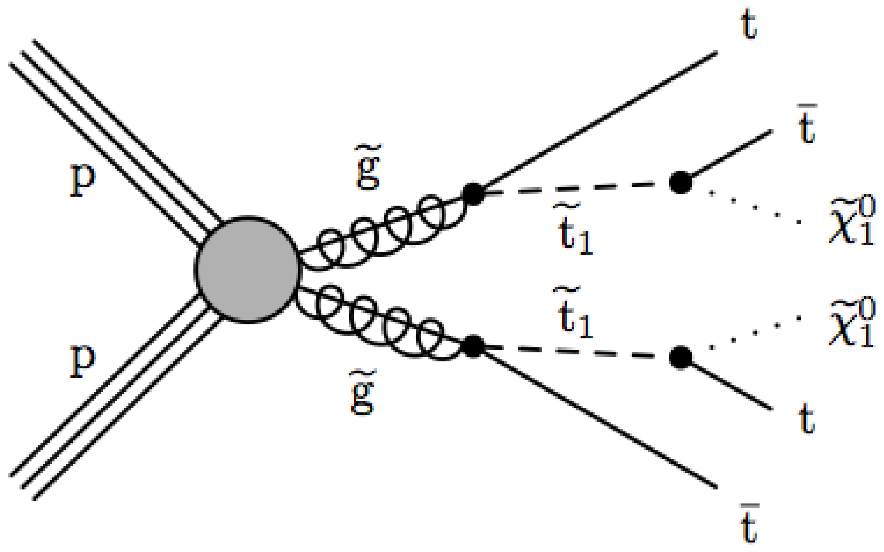
A gluino pair is produced in the *pp* interaction. Each gluino decays to a top quark and a top squark. The latter decays to a top quark and the neutralino.



Simplified models and boosted objects

Simplified models (SMSs) are effective models with the minimal particle content necessary to produce SUSY-like final states contributing to the channels of interest.

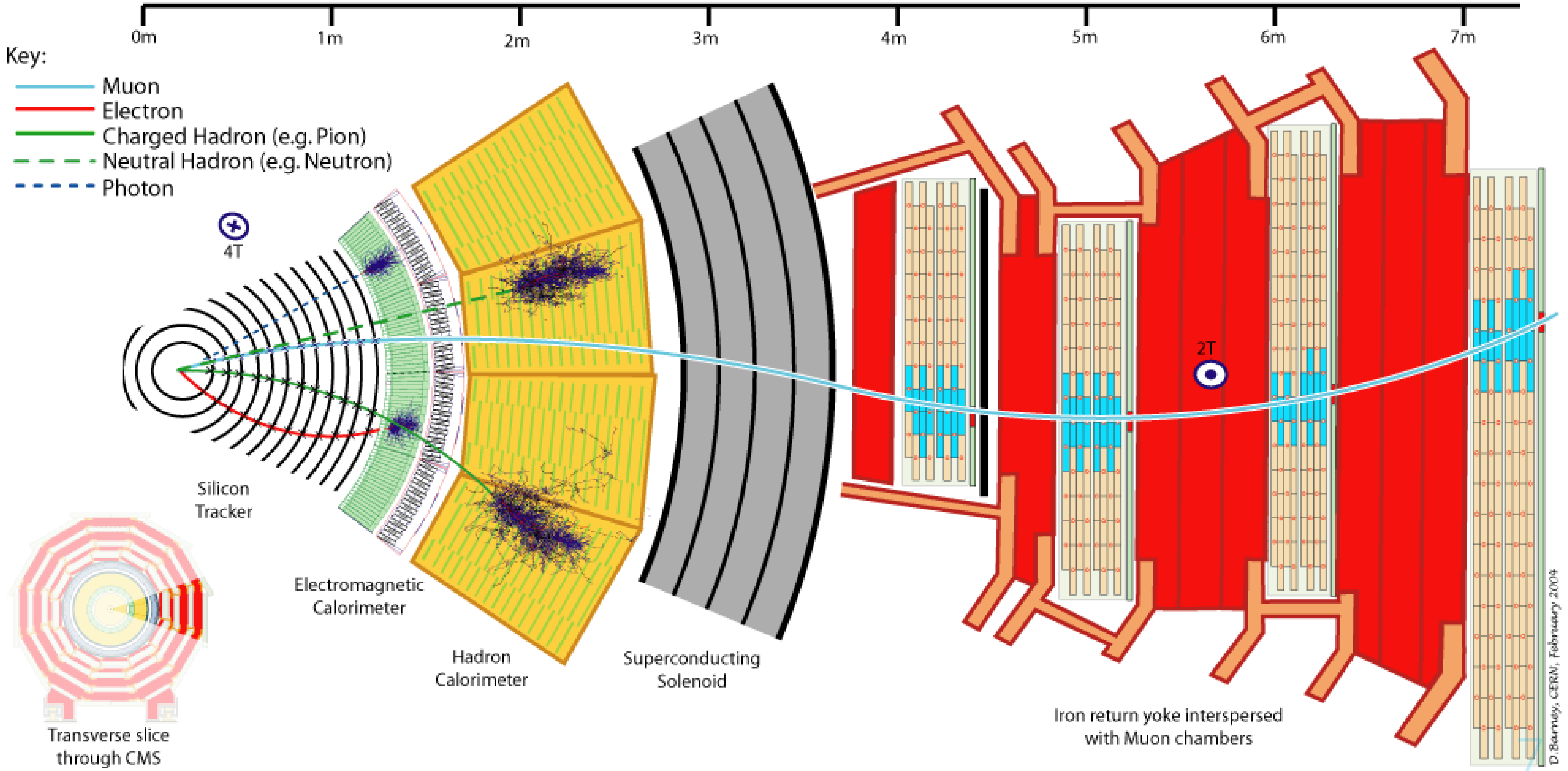
A gluino pair is produced in the *pp* interaction. Each gluino decays to a top quark and a top squark. The latter decays to a top quark and the neutralino.



$m_{\tilde{g}} > 1$ TeV and $m_{\tilde{t}} \approx O(100)$ GeV.

This leads to a significant mass gap between the two particles and thus the top quark from the gluino decay will receive a large boost.

The detector: Compact Muon Solenoid



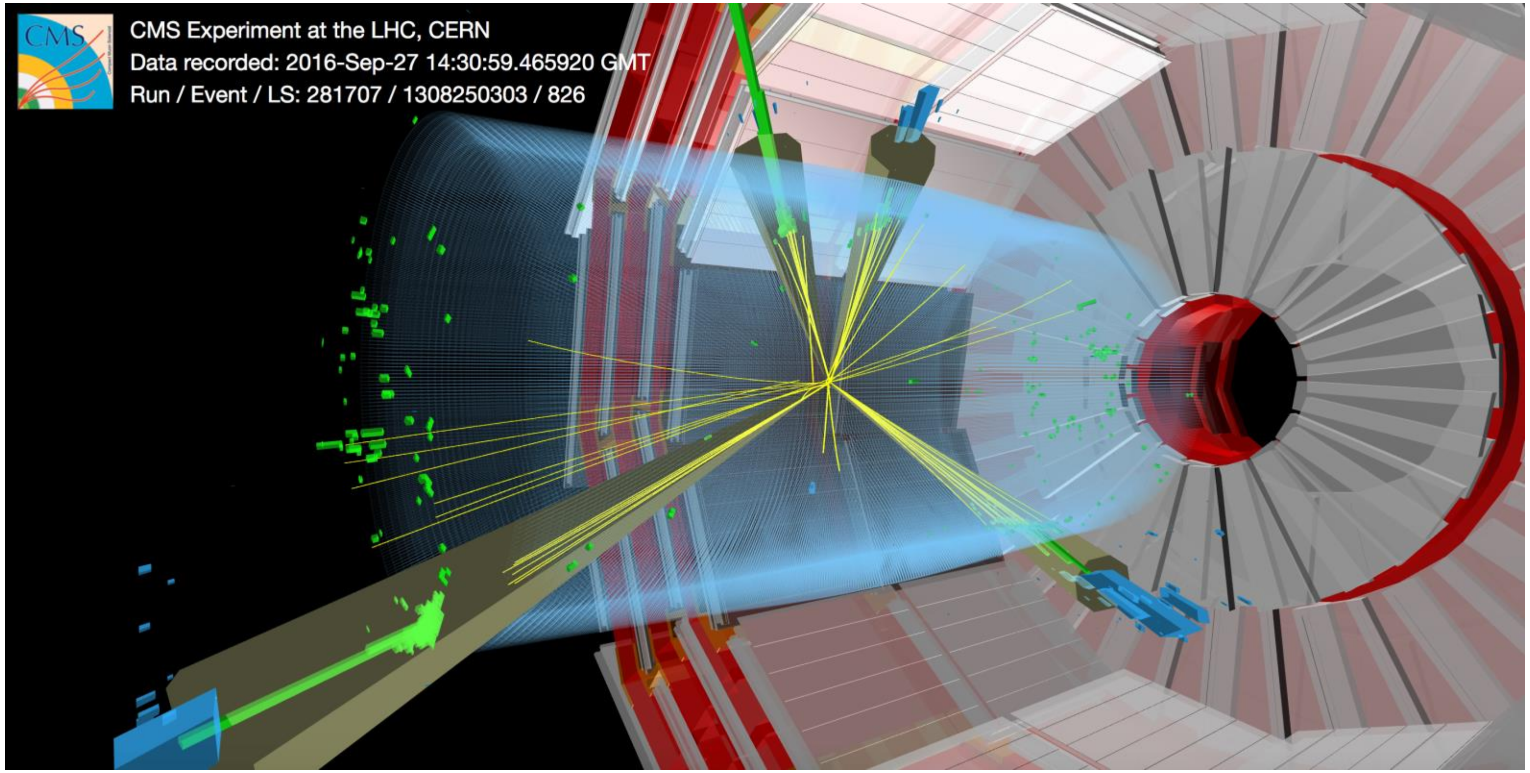
Main observables

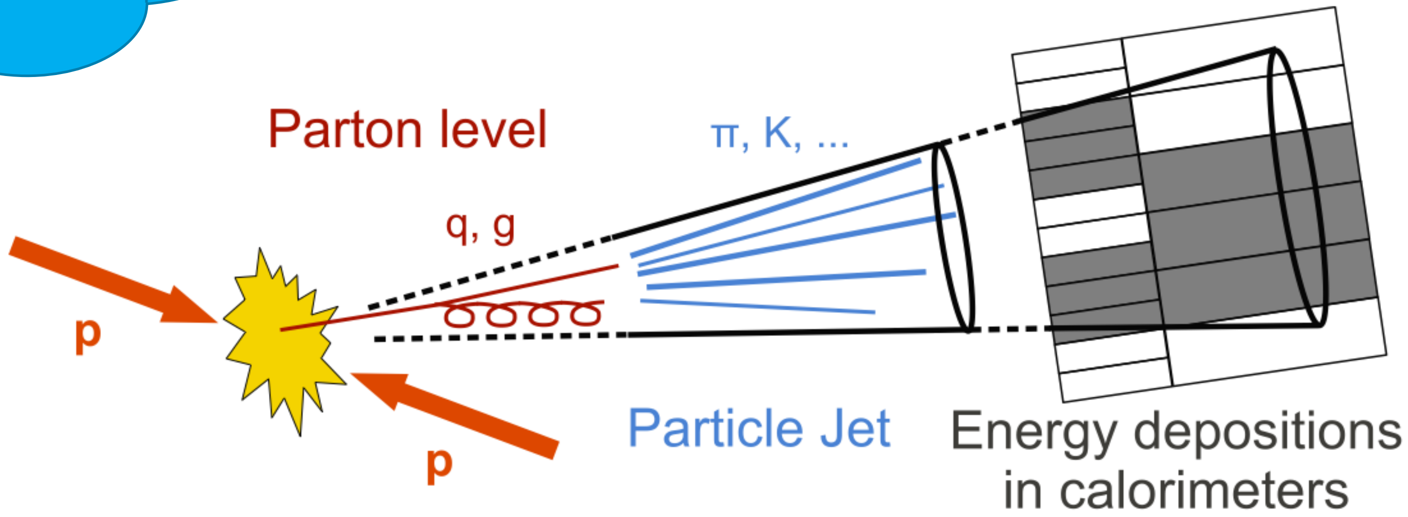
- $\vec{p}_T$
- Number of (b-)jets

Main observables

- $\vec{p}_T$
- Number of (b-)jets

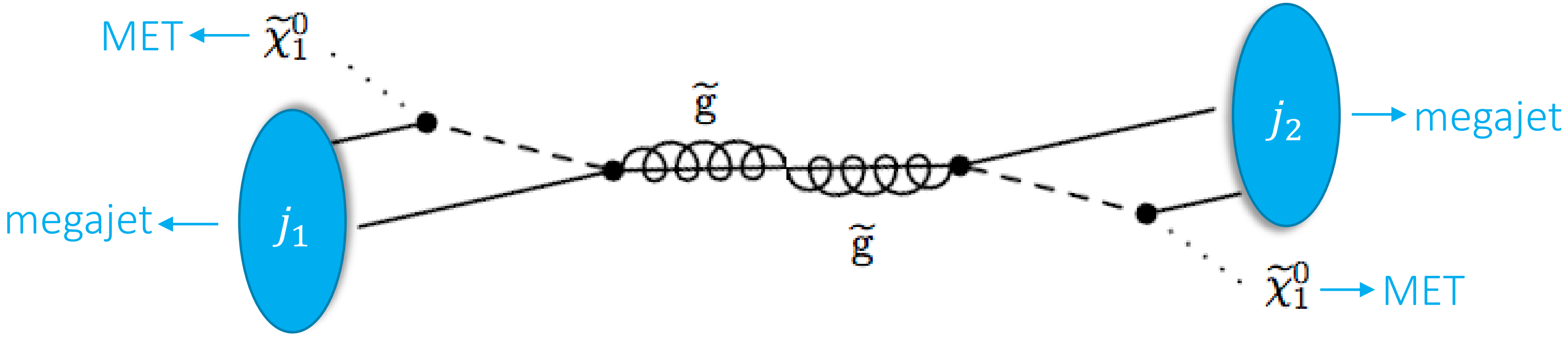
A jet is a number of particles originating from a single point, propagating in almost the same direction, and forming a cone.





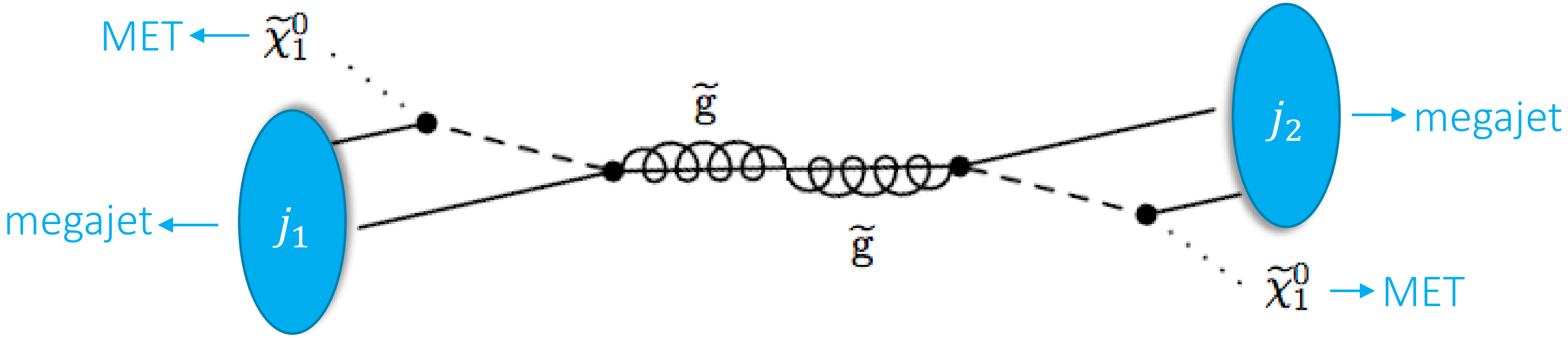
Main observables

- $\vec{p}_T$
- Number of (b-)jets
- MET = $| -\sum_{all} \vec{p}_T |$
- $H_T = \sum_{jets} |\vec{p}_T|$



Main observables

- $\vec{p}_T$
- Number of (b-)jets
- MET = $| -\sum_{all} \vec{p}_T |$
- $H_T = \sum_{jets} |\vec{p}_T|$



The search is performed in bins of the razor variables M_R and R^2 .

$$M_R = \sqrt{(|p^{j_1}| + |p^{j_2}|)^2 - \left(p_z^{j_1} + p_z^{j_2}\right)^2}$$

$$M_T^R = \sqrt{\frac{E_T^{miss}\left(p_T^{j_1} + p_T^{j_2}\right) - \vec{p}_T^{\,miss}\,(\vec{p}_T^{\,j_1} + \vec{p}_T^{\,j_2})}{2}}$$

$$R = \frac{M_T^R}{M_R}$$

M_R is like H_T and

M_T^R is like MET

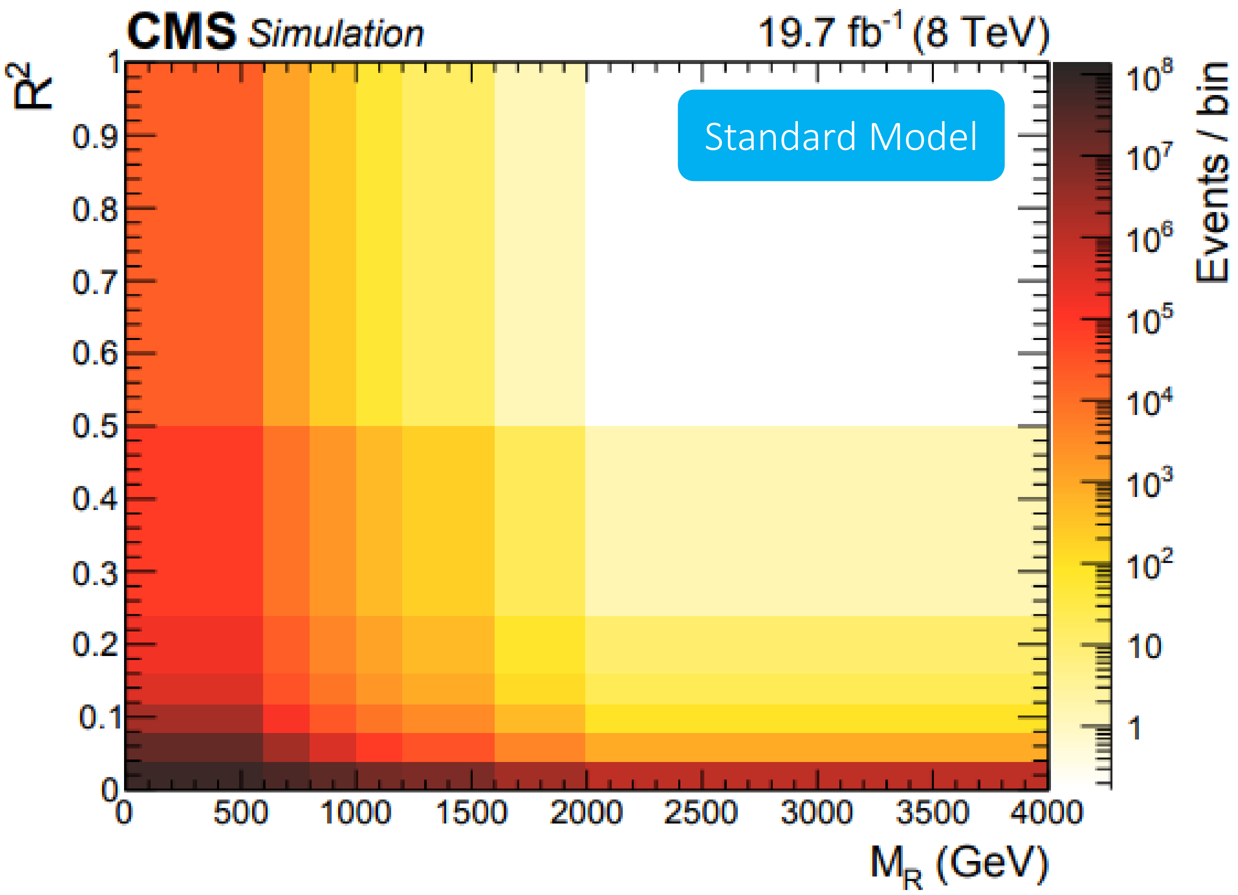
R then is the ratio of the invisible and visible energy in the event.

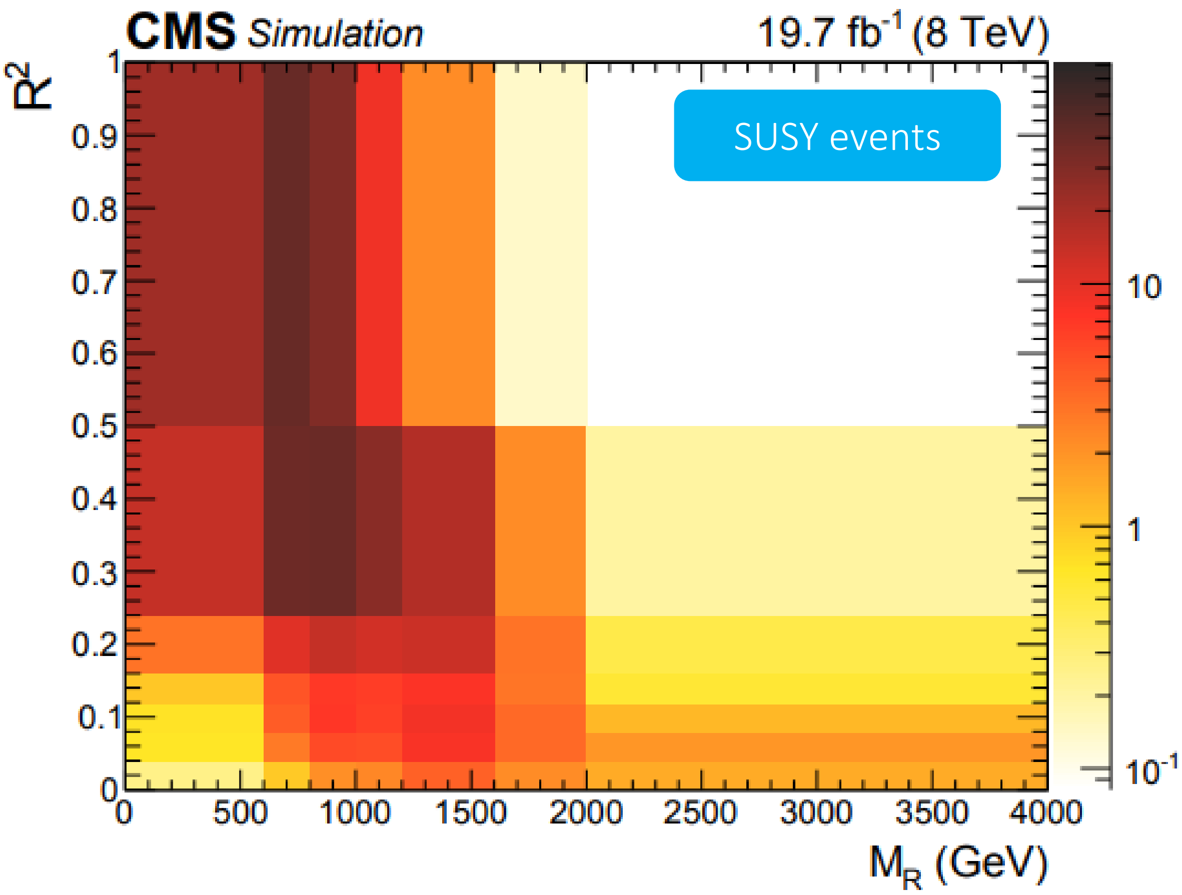
Main observables

How good discriminators are the Razor variables?

Main observables

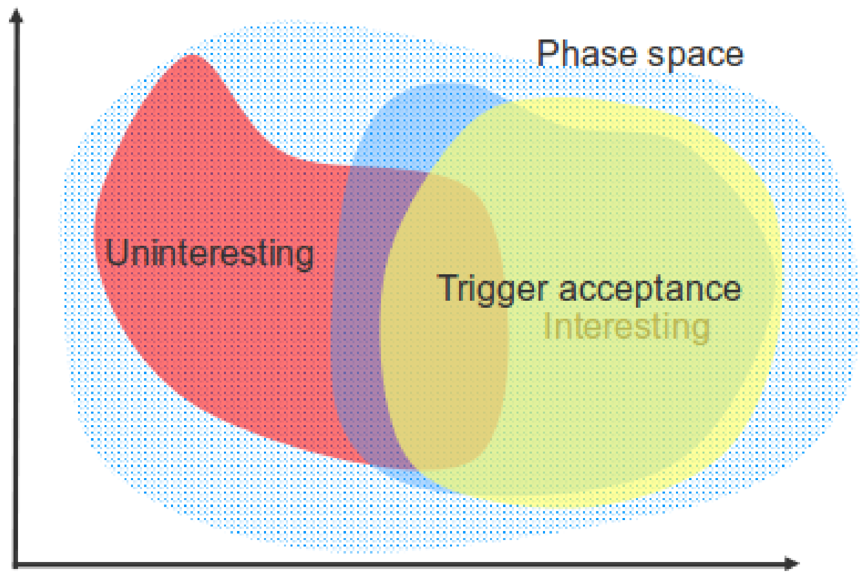
How good discriminators are the Razor variables?



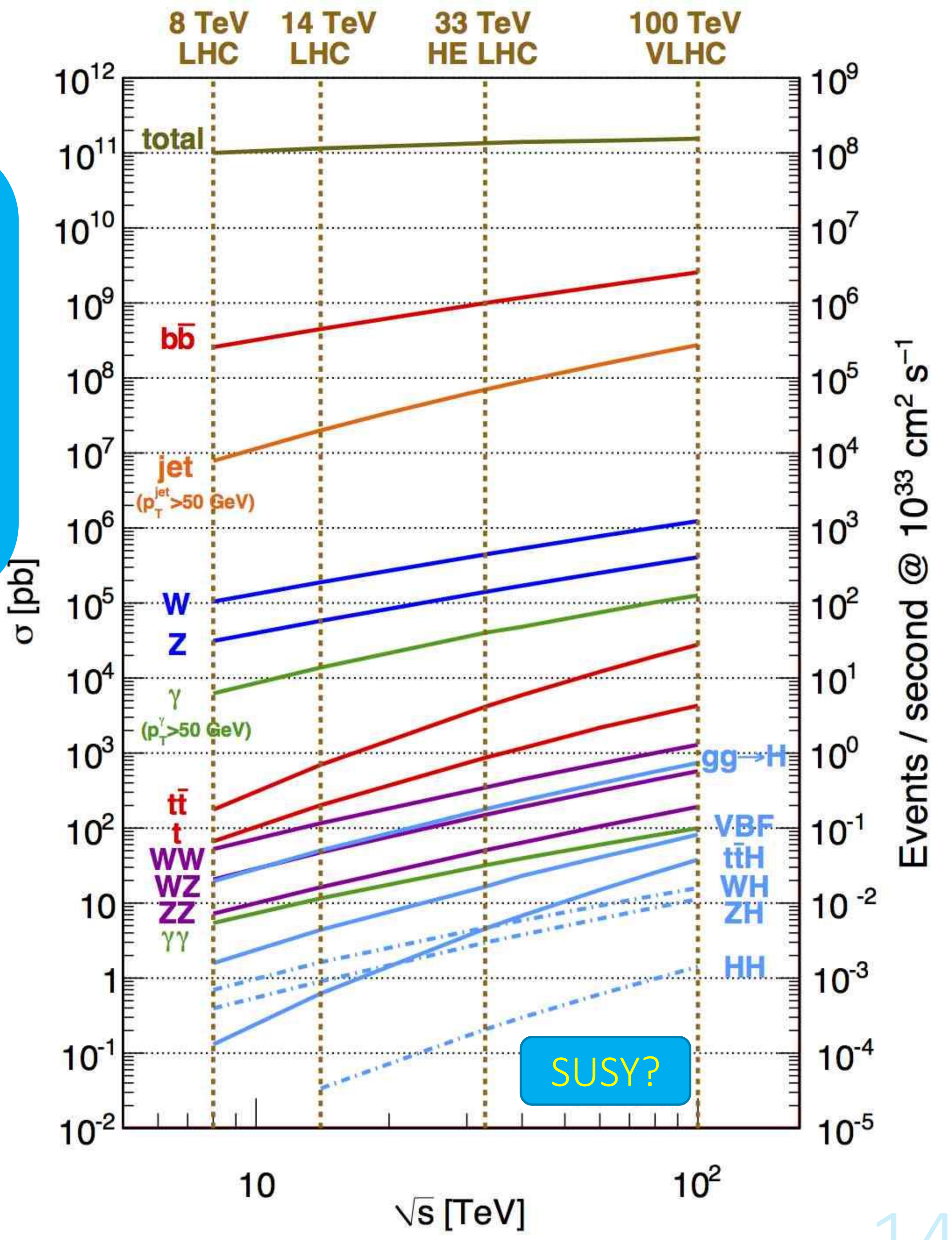


Trigger efficiencies

- The interesting events (signal) overlap with the uninteresting events (background) in the phase space spanned by the discriminating variables.
- The idea is to find triggers that select as much of the signal events as possible. But some background is unavoidable.



Courtesy of R. Jacobson (CERN).

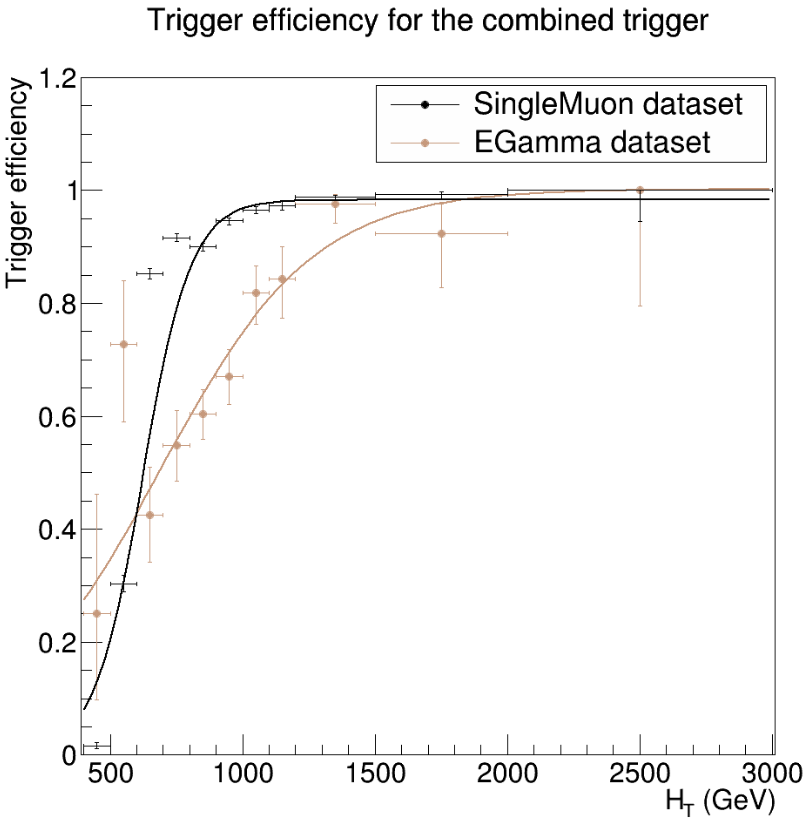


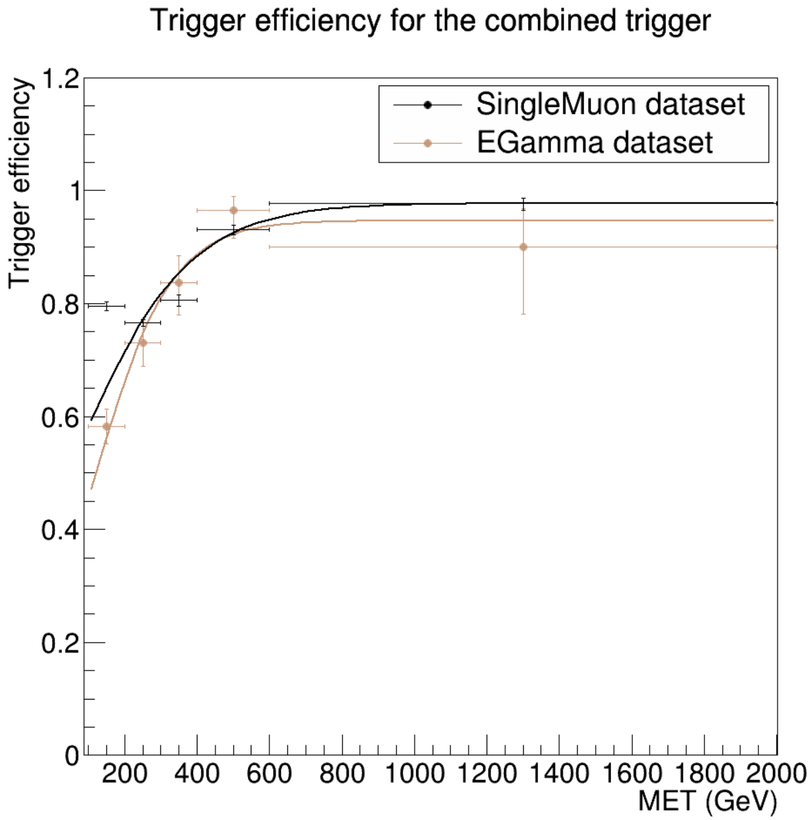
Trigger efficiencies

The idea is to find triggers that select as much of the signal events as possible.

First, 5 trigger families were studied, then choosing the best one, the triggers in that category (6 triggers) were checked. Here I only show the final plots (and only with two variables). More plots (13 pcs) are available here.

$$\epsilon_{trigger} = \frac{\text{baseline and preselections \&\& control trigger \&\& trigger under study}}{\text{baseline and preselections \&\& control trigger}}$$





Data/MC comparison

Events are categorized according to the number of leptons, jets, and b-tagged jets identified.

6 control regions (categories) were created:

- QCD enriched,
- Top enriched,
- Photon enriched,
- W enriched,
- 1-lepton
- Z enriched

Data/MC comparison

Events are categorized according to the number of leptons, jets, and b-tagged jets identified.

6 control regions (categories) were created:

- QCD enriched,
- Top enriched,
- Photon enriched,
- W enriched,
- 1-lepton
- Z enriched

To compare with data, the MC is reweighted with w.

$$w = \frac{\left(\int \mathcal{L}\, dt\right) \cdot \sigma(process)}{N_{events}} \; \epsilon_{trigger}$$

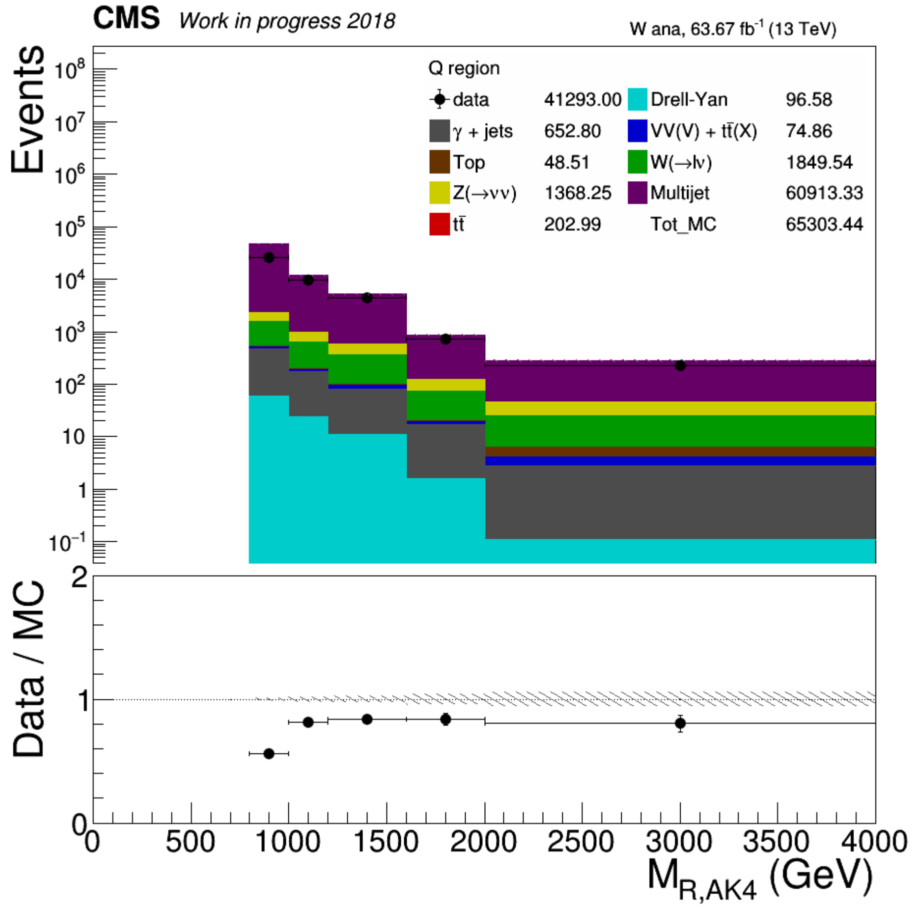
In 2018 $\int \mathcal{L}\, dt$ (i.e. collected data) was 63.67 fb^{-1}

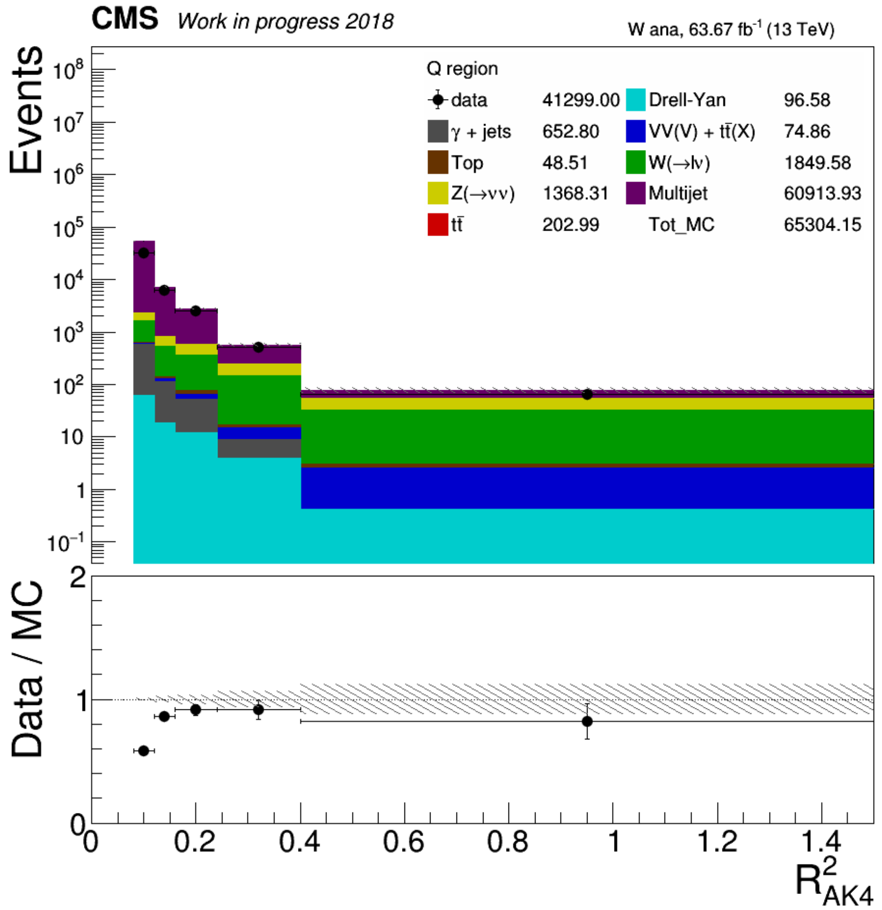
Cross sections (σ) for 2018 were calculated using the CMS software.

Trigger efficiency was measured (previous slide).

Data/MC comparison

In all the 6 categories, the coordinate dependence, the p_T, the MET, HT, M_R and R^2 dependence was compared. I only show the M_R and R^2 now for the QCD region. The rest 6*7 plots can be checked here.





Background estimation

The estimated number of events in the signal cross-check region for the multijet, W(→lν)+jets, and top quark processes is computed as follows:

$$N^{S'}_{multijet} = \left(N^{Q}_{obs} - N^{Q}_{other,MC}\right) / \left(\frac{N^{Q}_{multijet}}{N^{S'}_{multijet}}\right)_{MC}$$

$$N^{S'}_{W(\to\ell\nu)} = \left(N^{W}_{obs} - N^{W}_{other,MC}\right) / \left(\frac{N^{W}_{W(\to\ell\nu)}}{N^{S'}_{W(\to\ell\nu)}}\right)_{MC}$$

$$N^{S'}_{TTJ+T} = \left(N^{T}_{obs} - N^{T}_{multijet} - N^{T}_{other,MC}\right) / \left(\frac{N^{T}_{TTJ+T}}{N^{S'}_{TTJ+T}}\right)_{MC}$$

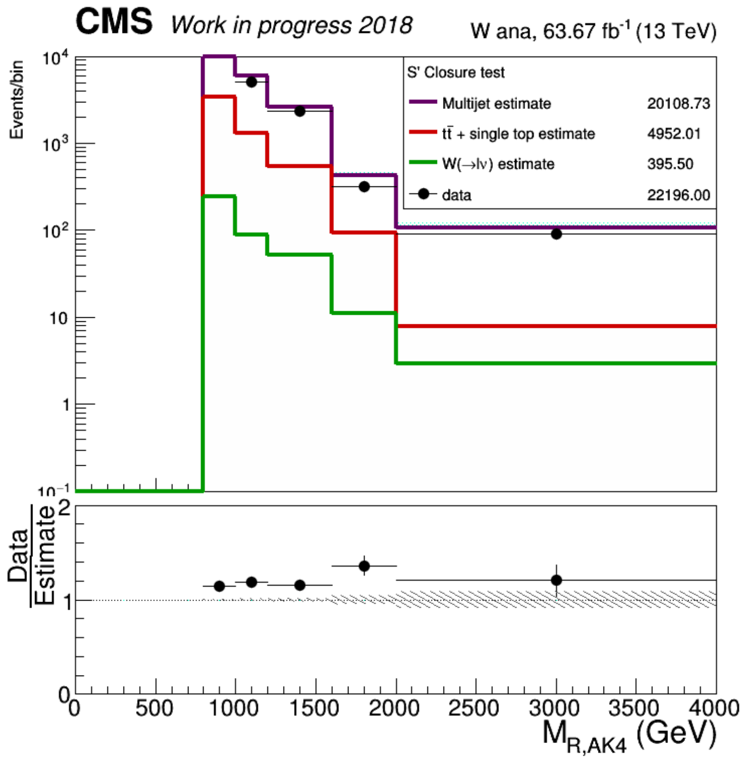
Background estimation

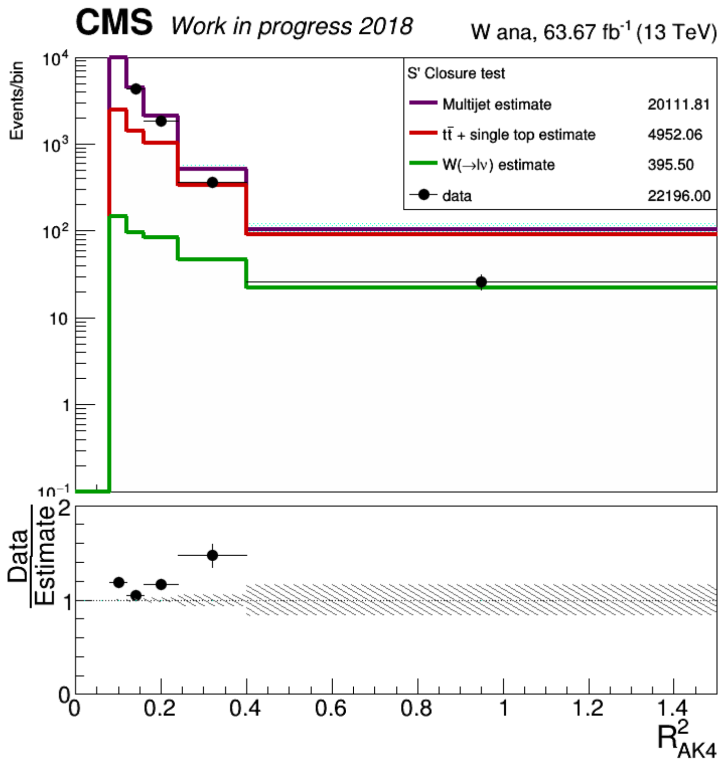
The estimated number of events in the signal cross-check region for the multijet, W(→lv)+jets, and top quark processes is computed as follows:

$$N^{S'}_{multijet} = \left(N^{Q}_{obs} - N^{Q}_{other,MC}\right) / \left(\frac{N^{Q}_{multijet}}{N^{S'}_{multijet}}\right)_{MC}$$

$$N^{S'}_{W(\to\ell\nu)} = \left(N^{W}_{obs} - N^{W}_{other,MC}\right) / \left(\frac{N^{W}_{W(\to\ell\nu)}}{N^{S'}_{W(\to\ell\nu)}}\right)_{MC}$$

$$N^{S'}_{TTJ+T} = \left(N^{T}_{obs} - N^{T}_{multijet} - N^{T}_{other,MC}\right) / \left(\frac{N^{T}_{TTJ+T}}{N^{S'}_{TTJ+T}}\right)_{MC}$$



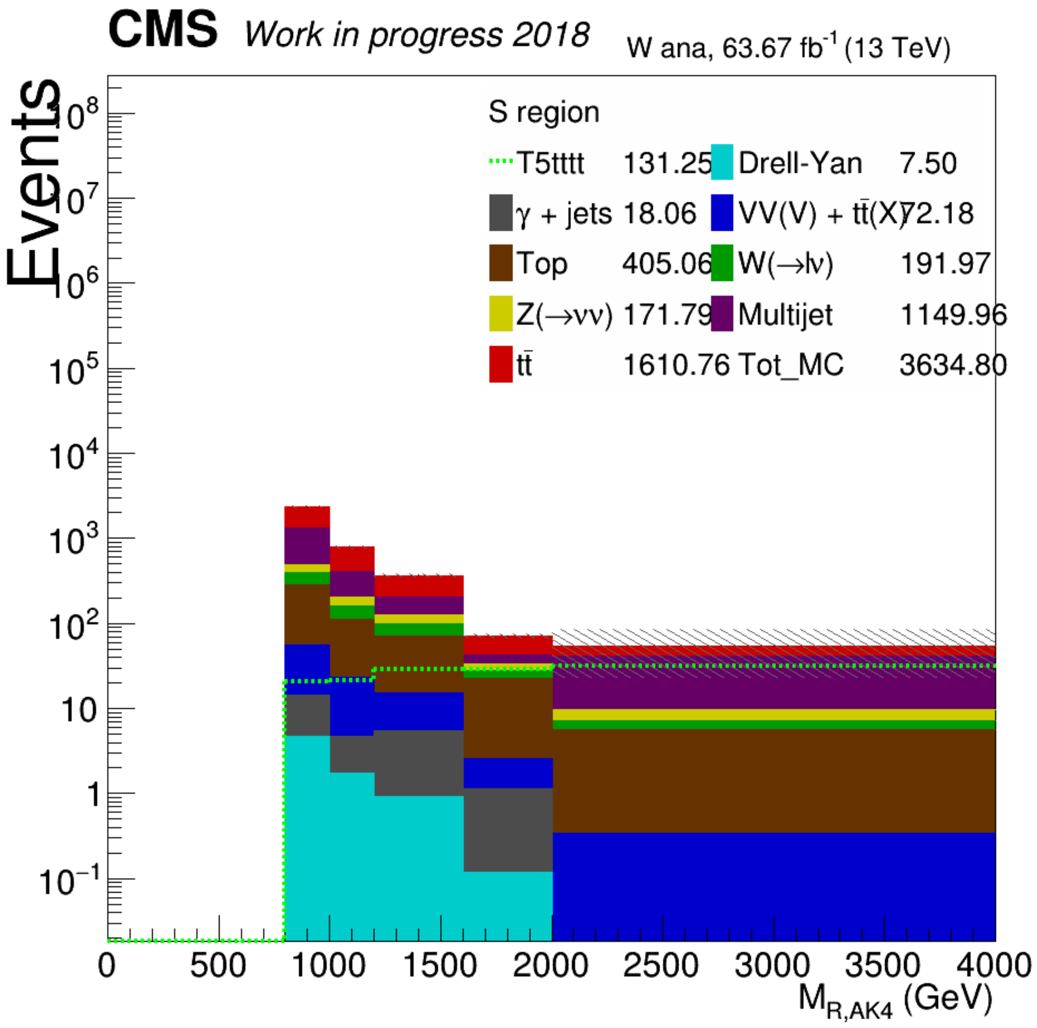


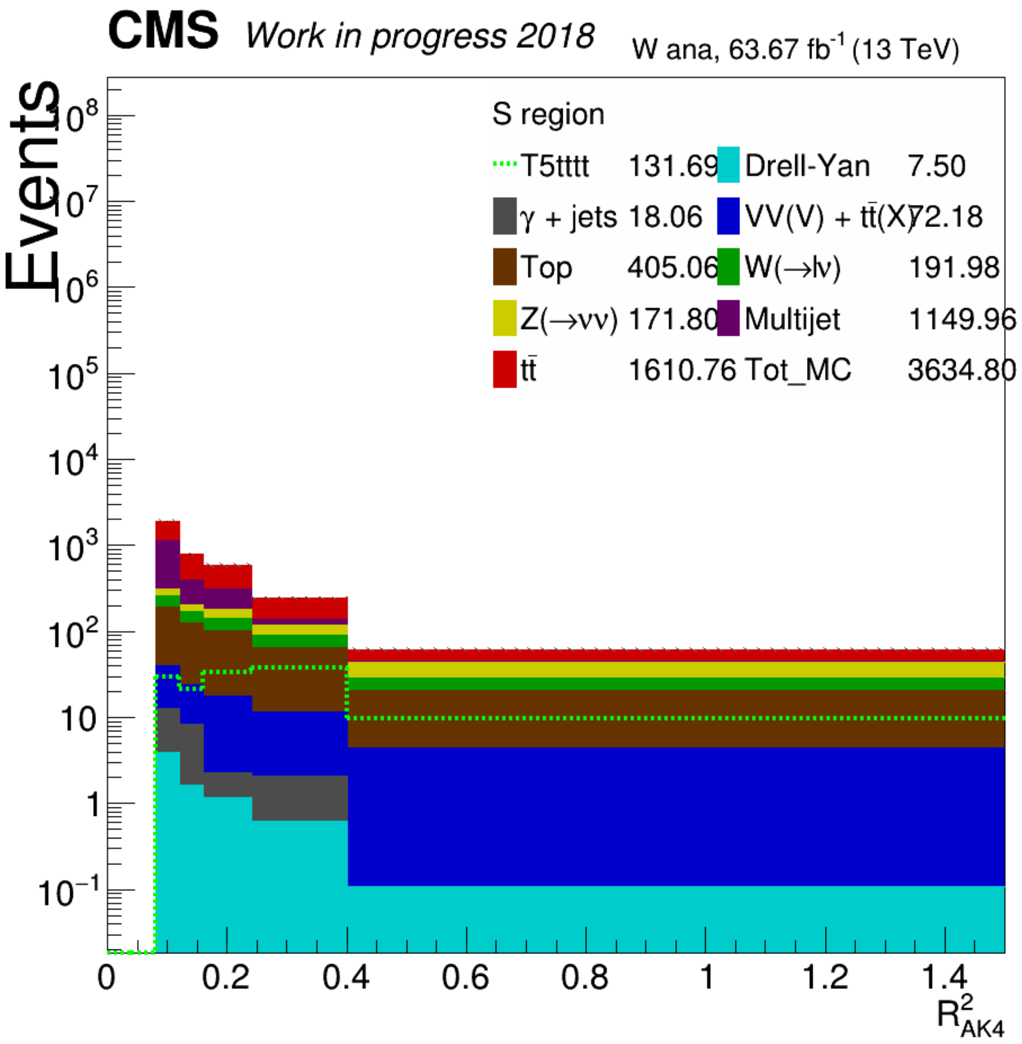
Again: similar plots exist for the QCD closure test as well.

Simulated signal process

The next step is to include the signal samples.

According to CMS policy, the data points in the signal region are not shown until approval.

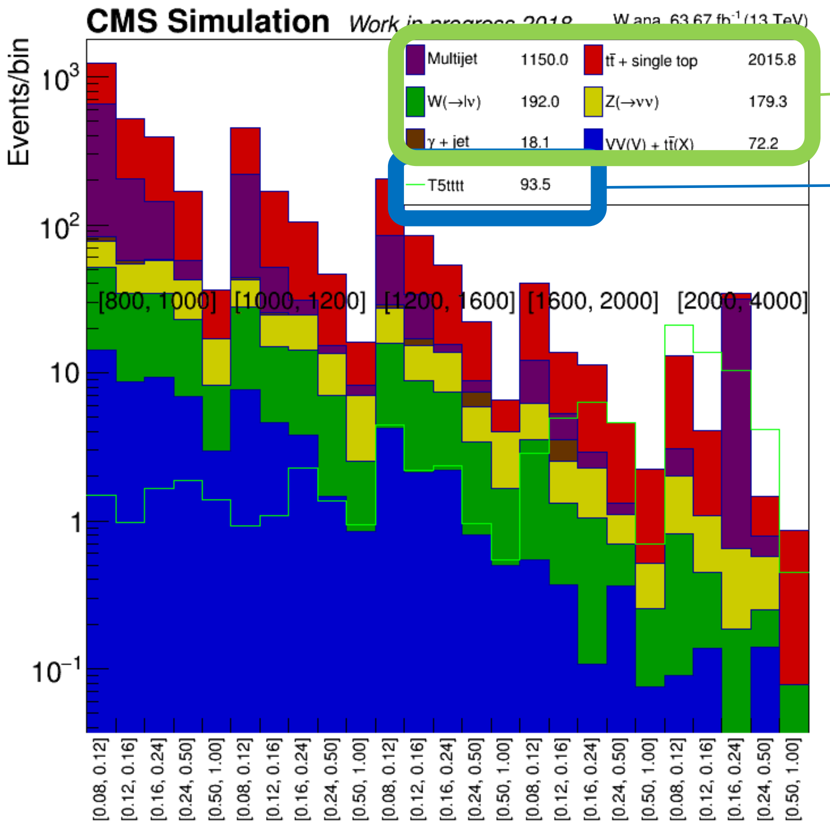




Other plots are here

Results and interpretation

The ratio of the signal events (S) and the square root of the signal+background (S+B) events is studied.



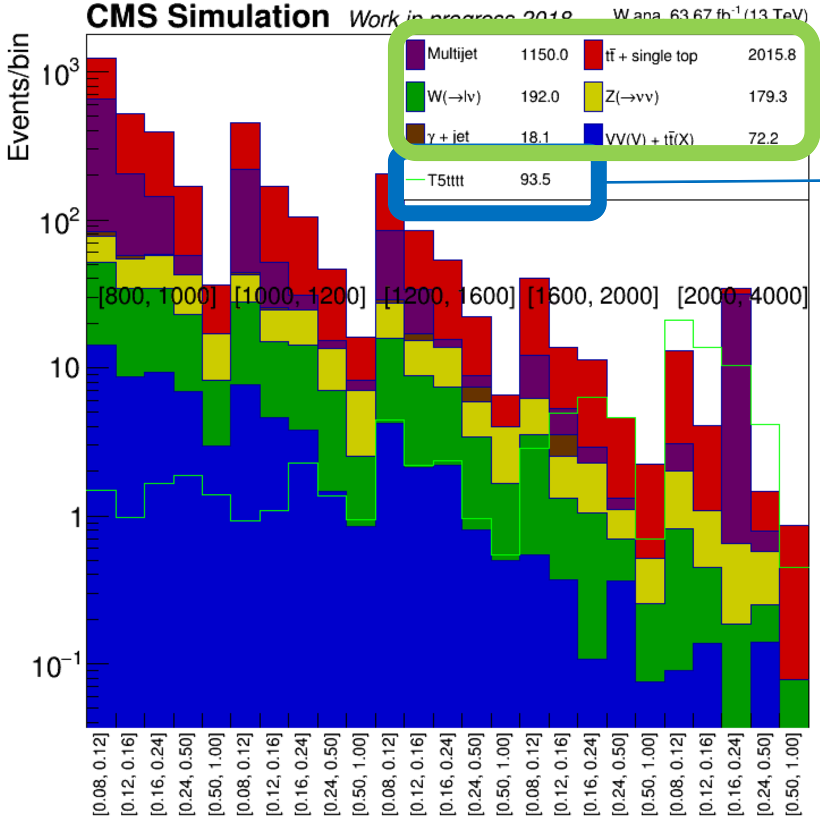
The sum is B

This is S

$S/\sqrt{S+B}$=1.53

Results and interpretation

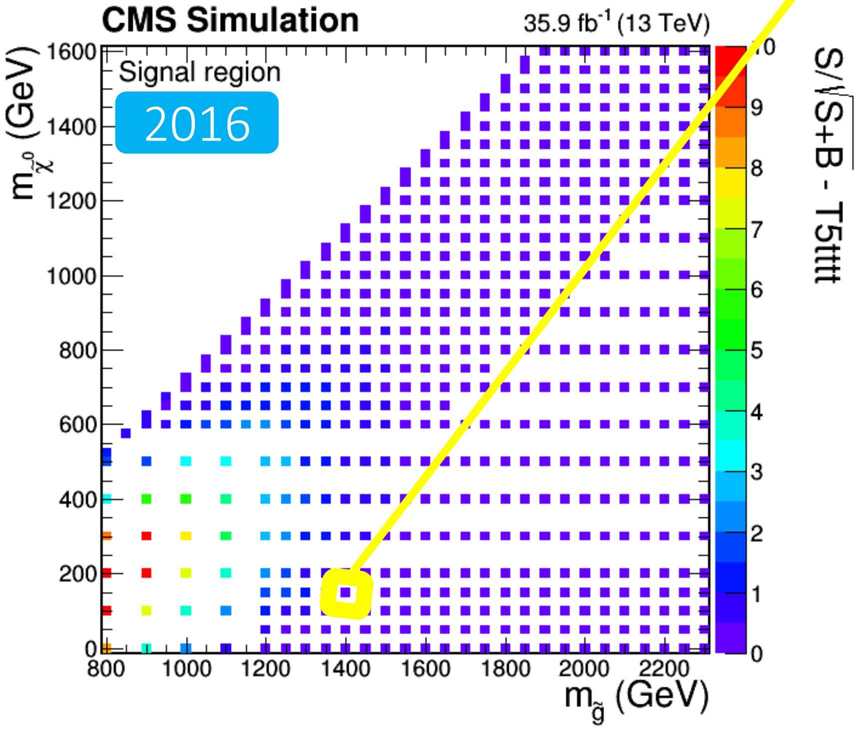
The ratio of the signal events (S) and the square root of the signal+background (S+B) events is studied.



The sum is B

This is S

$S/\sqrt{S+B}$=1.53 <=> $S/\sqrt{S+B}$= 0.88



Results and interpretation

Comparing this value to the value in 2018 data, we see that there is a factor of 1.73 multiplication.

The collected data increased from 35.9 fb^{-1} to 63.67 fb^{-1} with a factor of 1.8. We would expect the sensitivity to improve with the square root of that, i.e. with a factor of 1.34.

The results show improvement beyond what we would expect from purely statistical improvement. This is due to the better tagging techniques used.

Conclusion

A search for SUSY using the razor variables and boosted object tagging techniques was presented based on CMS data from 2018 with 63.67 fb^{-1} of *pp* collisions at $\sqrt{s}$ = 13 TeV.

The observed event yields in the cross-check region are found to be in line with the predicted backgrounds from SM processes.

- Consequently, no evidence of a signal is expected.

Sensitivity of the signal shows improvement beyond what we would expect from purely statistical improvement.

Next steps

Another round of central reprocessing is expected.

After that is done, the next steps are to

- collect the systematic errors,
- generate the SUSY signal to every parameter point,
- calculate limits for the mass of the gluino and the neutralino,
- show the analysis to CMS
- and after the approval, unblind the data.

Afterwards, if no significant excess is found, observed limit will be set.

A publication is on its way, we should have a paper in about a year from now.
I have a conference proceedings about the 2016 data (arXiv:1904.12148)

Backup

What have I done in the research?

I carried out a full analysis on trigger efficiency with several triggers for both 2017 and 2018 data.

I compared MC Simulation to 2018 data in control regions.

- For this I calculated the cross section for the MC samples used.
- I categorized events into groups and compared data to MC.

I estimated the background for the signal region using the MC methods.

I processed the SMS signal sample shown earlier, calculated the event weights and the corresponding cross sections.

I calculated the search sensitivity for a point ($m_{\tilde{g}}$= 1400 GeV and $m_{\chi_1^0}$ = 300 GeV) using the signal sample.

Backup – Selections

R-parity: $(-1)^{3(B-L)+2s}$

Baseline selection:

- the event is removed if the calorimeter did not function correctly due to dead cells or noise,
- beam halo events that contain machine-induced particles flying with the beam, at large radius (up to 5 m) are removed,
- only those vertices are used that were flagged good based on the number of degrees of freedom and geometrical constrains,
- and PF muons that do not qualify the quality criterion are removed as well

Further selection:

• at least one AK8 jet, • at least 3 jets, • MR > 800 GeV and R2 > 0.08.

Backup – Jet clustering

Jets, in general, are reconstructed using sequential recombination algorithms which work in the following way:

1. Define a distant parameter d_{ij} between entities i and j and d_{iB} between entity i and the beam.
2. Identify the smallest of the distances.
   (a) If it is d_{ij}, combine entities i and j.
   (b) If it is d_{iB}, call i a jet and remove it from the list of entities.
3. Recalculate the distances and repeat the procedure until no entities are left.

The difference in the algorithms is in the definitions of distances:

$$d_{ij} = min\left(k_{T,i}^{2p}, k_{T,j}^{2p}\right) \frac{(y_i - y_j)^2 + (\phi_j - \phi_i)^2}{R^2} \tag{8}$$

and

$$d_{iB} = k_{T,i}^{2p} \tag{9}$$

where $k_{T,i}$, y_i, ϕ_i, R are respectively the transverse momentum, rapidity, azimuth of particle i and the radius parameter.

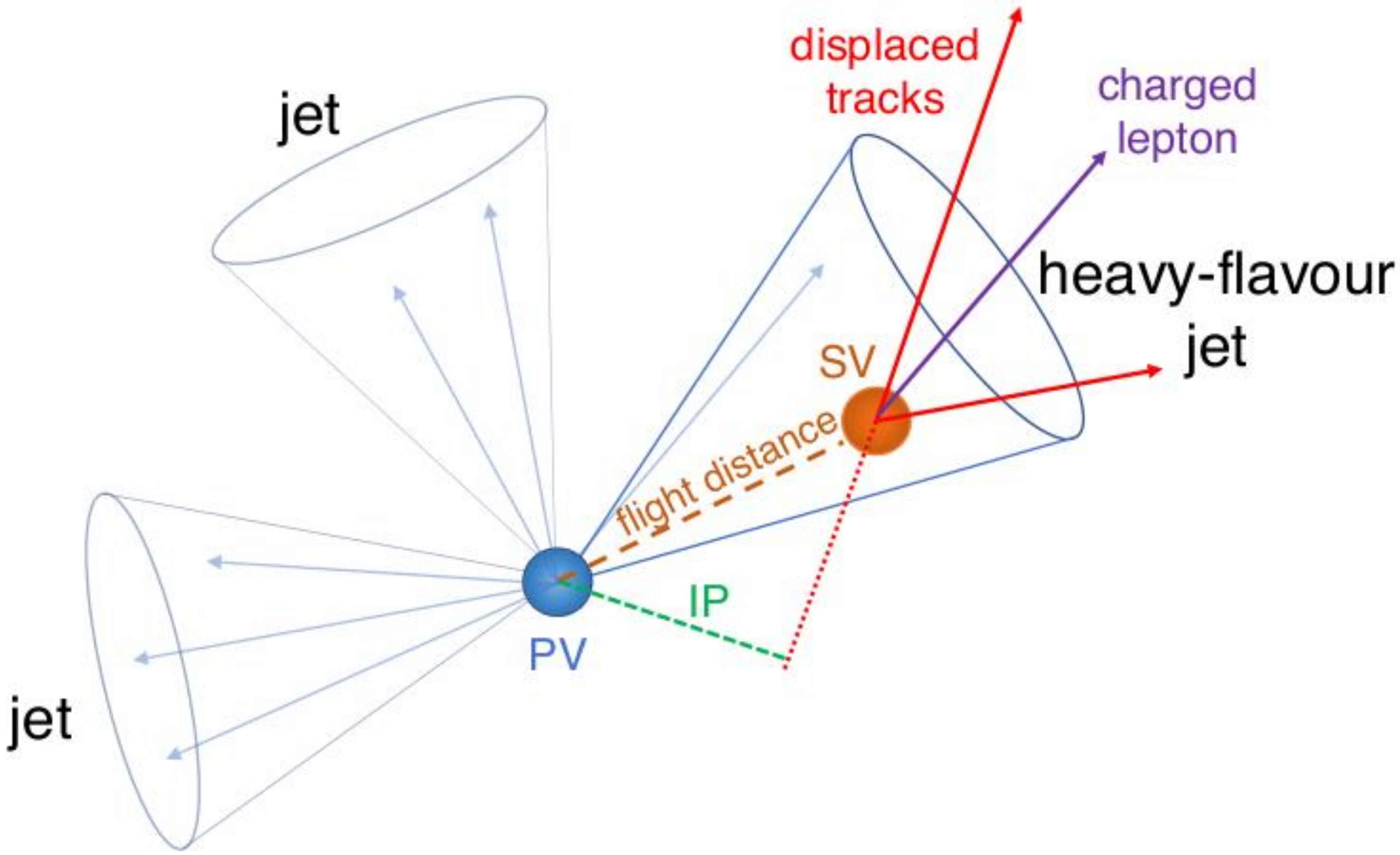
Backup – Trigger results

Trigger name	Number of events
All events:	42,651,472
HLT_AK8PFJet450:	144,779
HLT_AK8PFJet400_TrimMass30:	1,403,728
HLT_AK8PFHT800_TrimMass50:	582,503
HLT_PFHT1050:	518,275
HLT_PFHT500_PFMET100_PFMHT100_IDTight:	2,204,026
HLT_IsoMu27	7,810,493

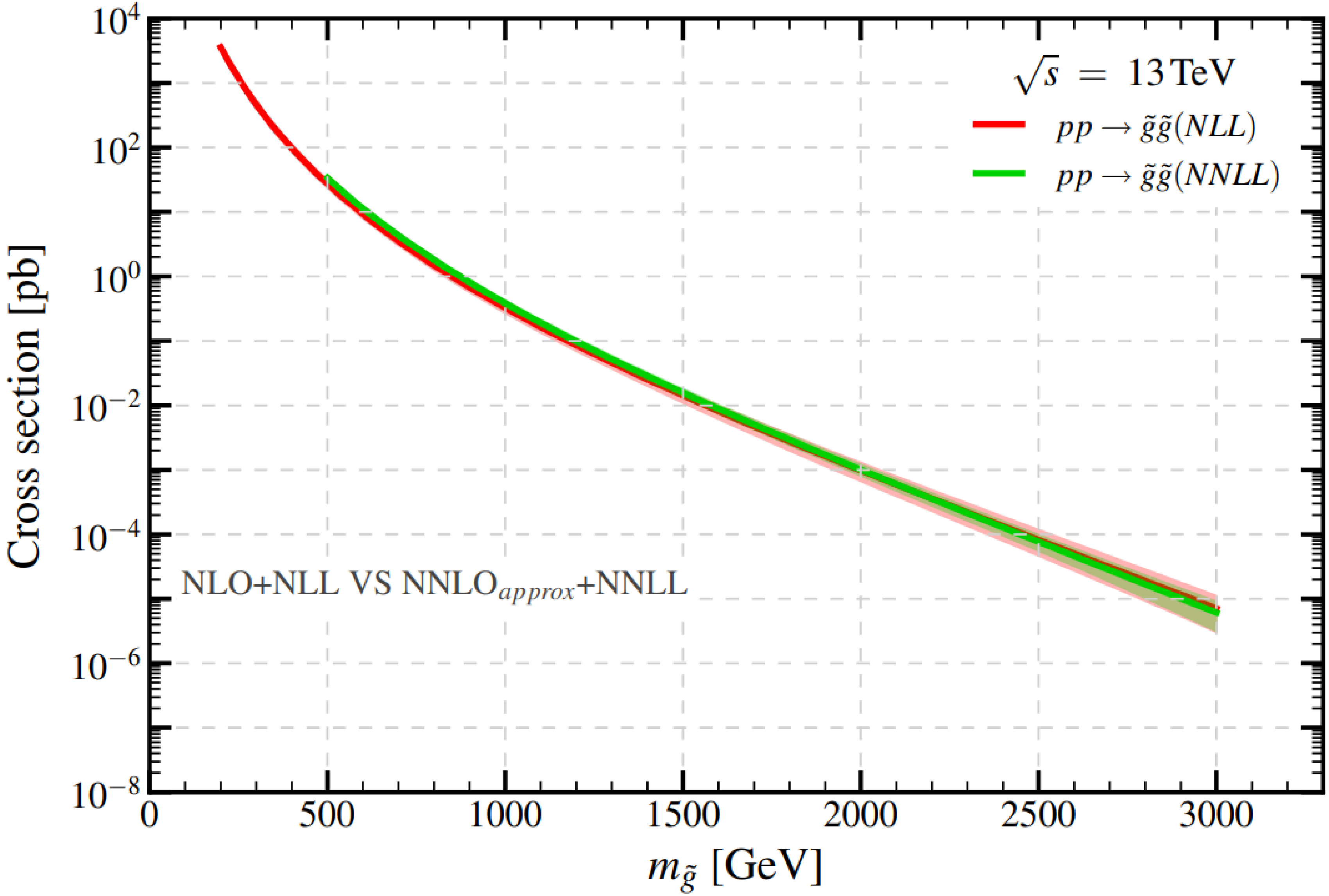
Backup – Region definitions

	S	S'	Q	Q'	G	T	W	L	Z
Preselection	✓	✓	✓	✓	✓	✓	✓	✓	✓
$n_{EleVeto}$	0	0	0	0	0	0 or 1	0 or 1	0 or 1	N/A
n_{MuVeto}	0	0	0	0	0	1 or 0	1 or 0	1 or 0	N/A
$n_{TauVeto}$	0	0	0	0	0	N/A	N/A	N/A	N/A
$n_{MediumBTag}$	≥1	≥1	N/A	N/A	N/A	≥1	N/A	N/A	N/A
$n_{LooseBTag}$	N/A	N/A	0	0	N/A	N/A	0	0	N/A
$n_{TightWTag}$	≥1	≥1	N/A	N/A	N/A	≥1	N/A	N/A	N/A
$n_{TightWAntiTag}$	N/A	N/A	≥1	≥1	N/A	N/A	N/A	N/A	N/A
$n_{WMassTag}$	N/A	N/A	N/A	N/A	≥1	N/A	≥1	≥1	≥1
$\Delta\phi_{Razor}$	<2.8	≥2.8	≥2.8	<2.8	<2.8	<2.8	<2.8	<2.8	<2.8
$n_{PhotonSelect}$	N/A	N/A	N/A	N/A	1	N/A	N/A	N/A	N/A
$30 \leq m_T < 100$	N/A	N/A	N/A	N/A	N/A	< 100	✓	✓	N/A
$\|(M_{ll} - M_Z) < 10\|$	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	✓

Backup – B tagging



Backup – SUSY cross sections



Backup – HEM15/16 issue