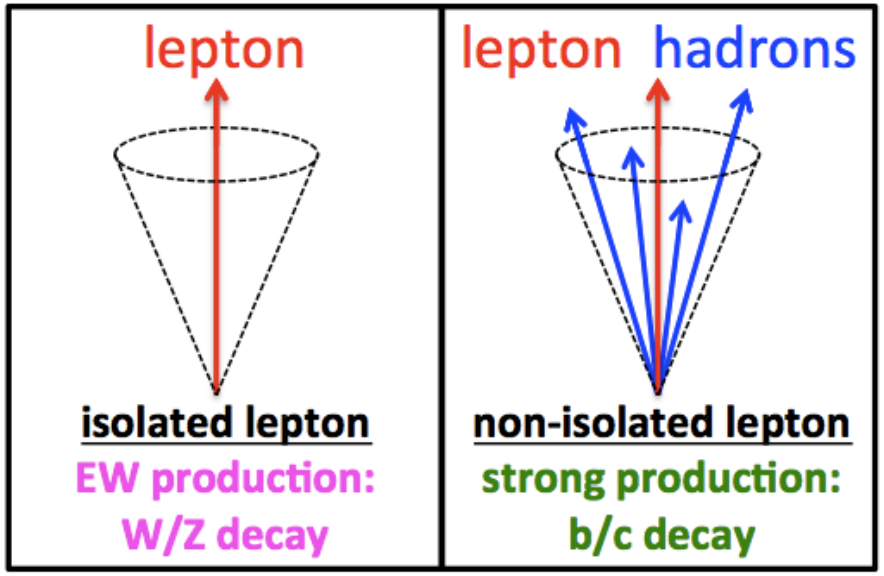
Search for New Physics with Same-Sign Isolated Dilepton Events with Jets and Missing Transverse Energy at CMS

Frank Golf,
on behalf of the CMS Collaboration

2011 Results: CMS-SUS-11-010
2010 Results: JHEP 1106 (2011) 077 and CMS-SUS-10-004

Physics Signature

- **Prompt, isolated same sign dileptons.**
  - Prompt lepton → from W or Z.
  - Isolated lepton → not in a jet.
  - Leptons from heavy flavor decays and decays in flight are considered background for this search. We refer to leptons that are not from a W/Z as "fake leptons".



- **Missing Transverse Energy (MET).**
- **Significant hadronic activity (Jets).**
  - Count number of jets and the scalar sum p_T (H_T) of these jets.

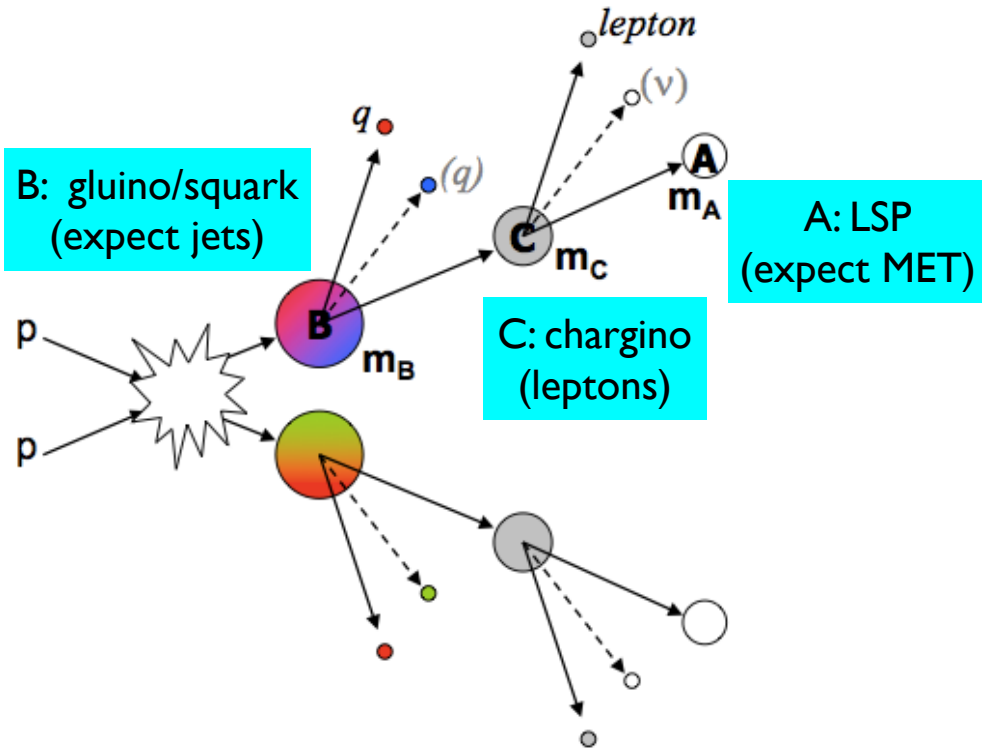
Motivation of Search Signature

Search is signature-based, not limited to a particular model

- Same sign dileptons are rare in the standard model → expect small backgrounds.
- Signatures with MET are motivated by evidence for dark matter.
- Many new physics (NP) scenarios with large cross sections are produced strongly → leptons accompanied by significant hadronic activity.
- Many NP scenarios lead to enhanced third generation production → search for taus.
- Sensitive to a variety of new physics models: SUSY, extra dimensions, same-sign top [JHEP 1108 (2011) 005], etc.

Searches

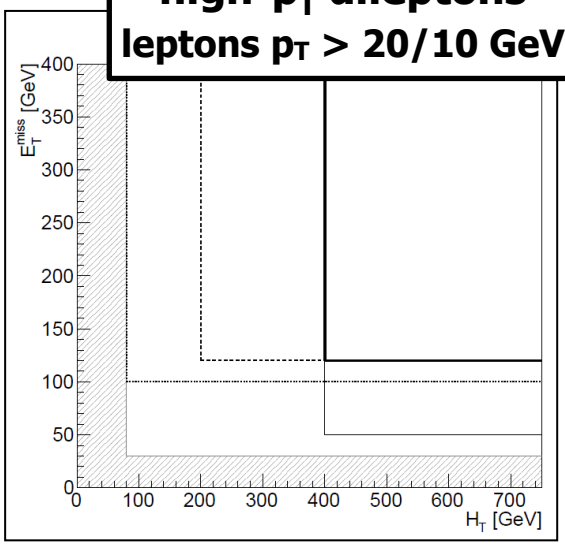
- Model independent search → cover as much phase space as possible.
- **Three searches:**
  - **High p_T dileptons (e, μ)**
    - Two different methods to estimate backgrounds (A1, A2).
  - **Inclusive dileptons (e, μ)**
    - Sensitive to low p_T leptons.
    - Two different methods to estimate backgrounds (A1, B).
  - **Final states with hadronic taus ($e\tau$, $\mu\tau$, $\tau\tau$)**
    - Sensitive to models w/ enhanced 3rd generation couplings.
    - Estimate backgrounds using methods similar to A1, A2.



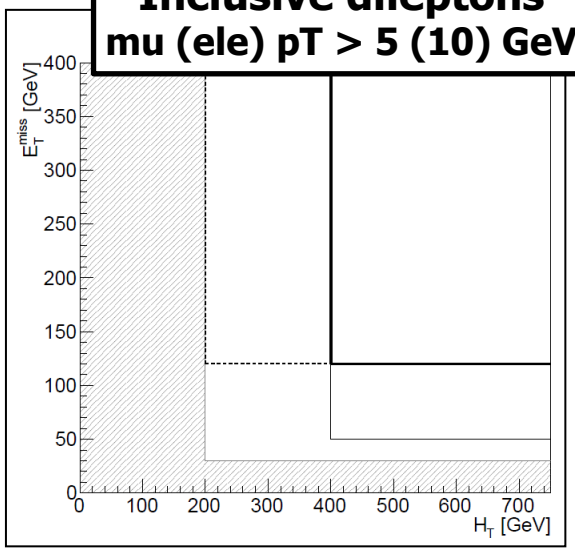
M_B drives cross section.
ΔM_{BC} drives H_T.
ΔM_{CA} drives lepton p_T.
ΔM_{BA} influences MET.

Search Regions

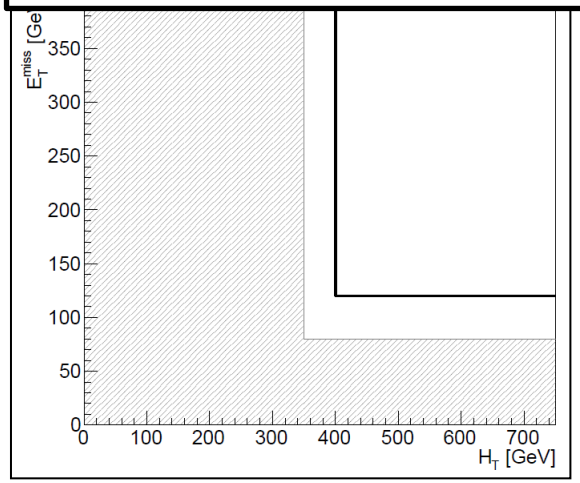
high-p_T dileptons
leptons p_T > 20/10 GeV



Inclusive dileptons
mu (ele) pT > 5 (10) GeV



τ_h dileptons
mu (ele) [tau] pT > 5 (10) [15] GeV



<u>Region 1</u> (ee, $\mu\mu$, $e\mu$) :
H_T> 400 GeV, MET > 120 GeV

<u>Region 2</u> (ee, $\mu\mu$, $e\mu$) :
H_T > 200 GeV, MET > 120 GeV

<u>Region 3</u> (ee, $\mu\mu$, $e\mu$) :
H_T > 400 GeV, MET > 50 GeV

<u>Region 4</u> (ee, $\mu\mu$, $e\mu$) :
H_T > 80 GeV, MET > 100 GeV

<u>Region 1</u> (ee, $\mu\mu$, $e\mu$) :
H_T > 400 GeV, MET > 120 GeV

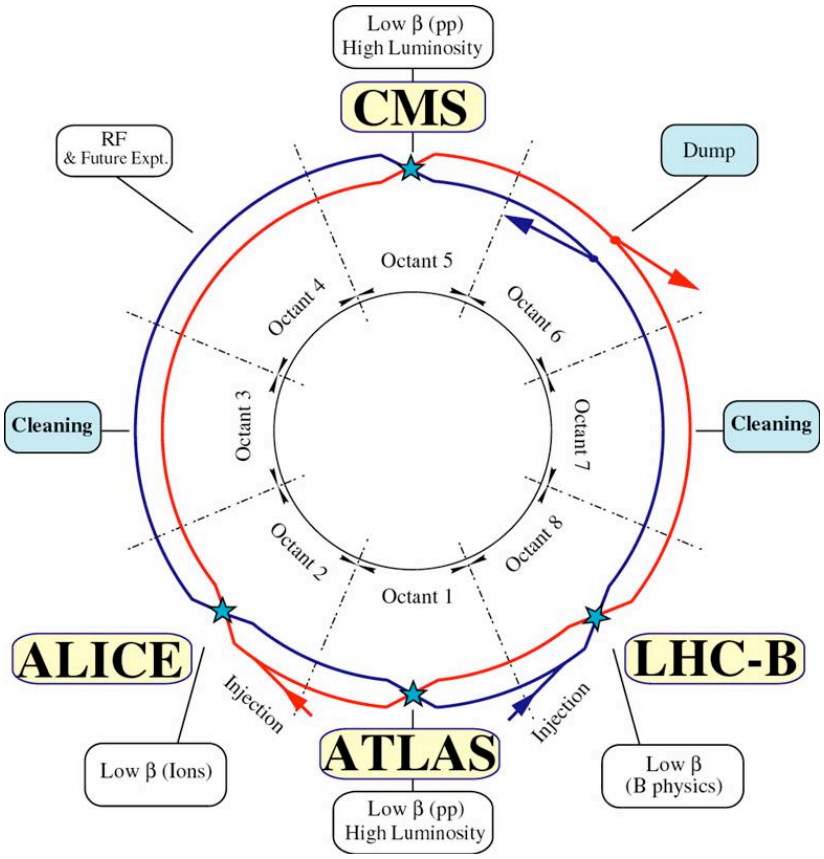
<u>Region 2</u> (ee, $\mu\mu$, $e\mu$) :
H_T > 200 GeV, MET > 120 GeV

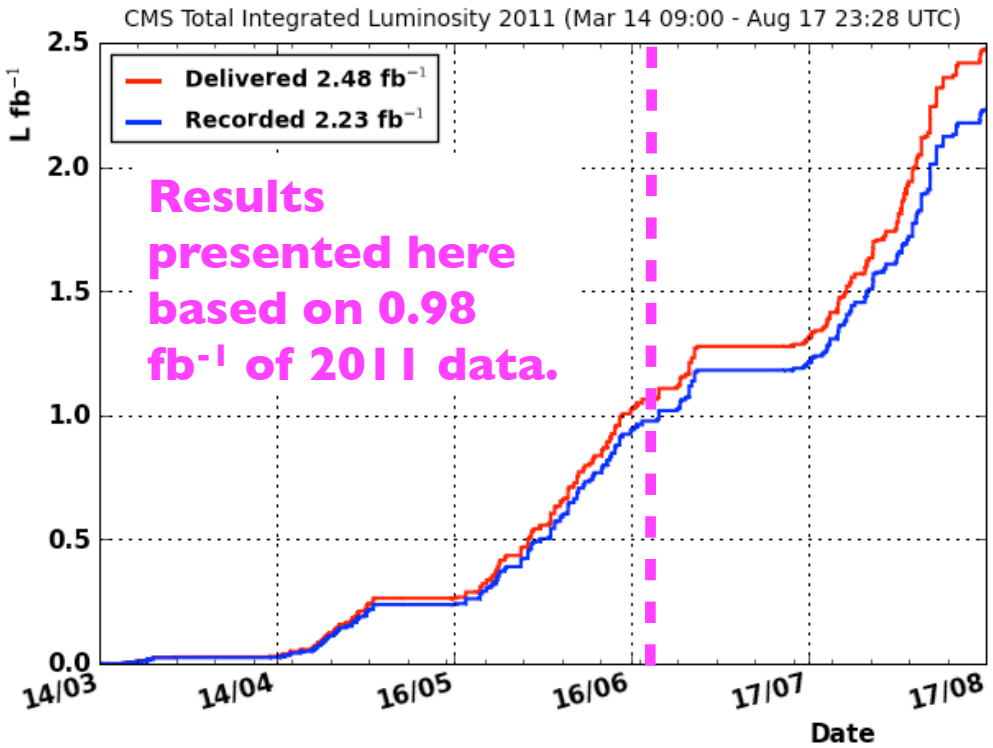
<u>Region 3</u> (ee, $\mu\mu$, $e\mu$) :
H_T > 400 GeV, MET > 50 GeV

<u>Region 1</u> ($e\tau$,$\mu\tau$, $\tau\tau$) :
H_T > 400 GeV, MET > 120 GeV

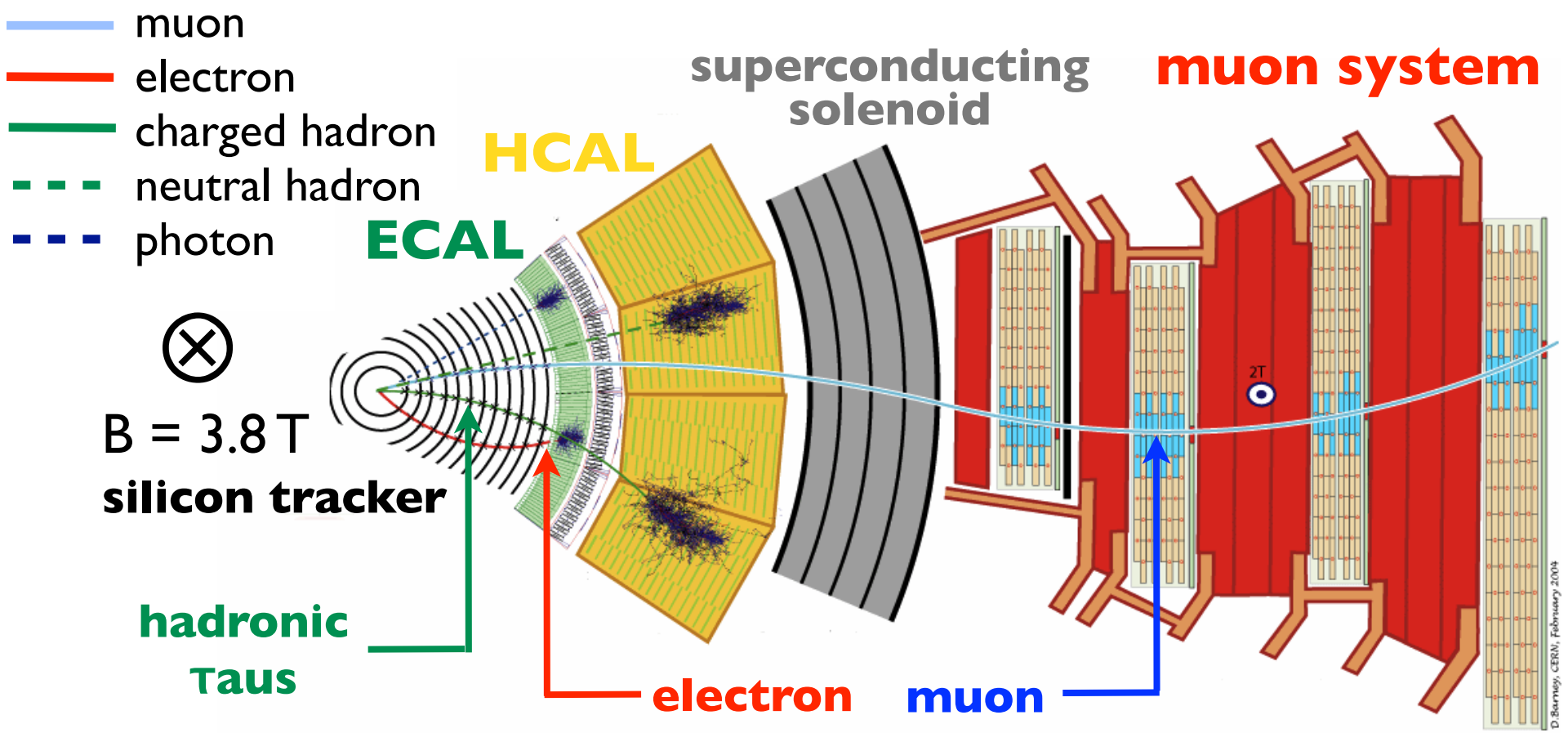
The LHC

- A 27 km circumference ring on the border of France and Switzerland.
- Produces proton-proton collisions at $\sqrt{s}$ = 7 Trillion eV (TeV).
- The data is collected by four experiments: **CMS**, ATLAS, LHCb, ALICE.





The CMS Detector



- **Electrons: shower in ECAL → deposit full energy**
- **Muons: minimum ionizing → pass through detector**
- **Taus: hadronic decay → 1 or 3 tracks from charged hadrons**

Background Estimation Strategy

- **Background from failures of lepton identification (fake leptons from jets).**
  - **Largest background to the same sign searches.**
  - Simulation indicates that the dominant contribution comes from ttbar events, where one lepton is from a *W* and the other lepton is a fake from a b-jet.
  - Estimate using data-driven methods.

- **Background from failures of charge reconstruction.**
  - **Small background to the same sign searches.**
  - Two opposite sign prompt isolated leptons with the sign of the charge of one of the leptons mis-measured.
  - Estimate using data-driven methods.

- **Background from SM sources of prompt, isolated same sign dileptons.**
  - Contribution from *WZ* and *ZZ* and from rare SM processes (ttbarW, single and double parton processes). Also, a small contribution from $W\gamma/Z\gamma$ (conversion not predicted by data-driven methods).
  - **Irreducible background can be significant → contributes 10 - 40% depending on the search region.**

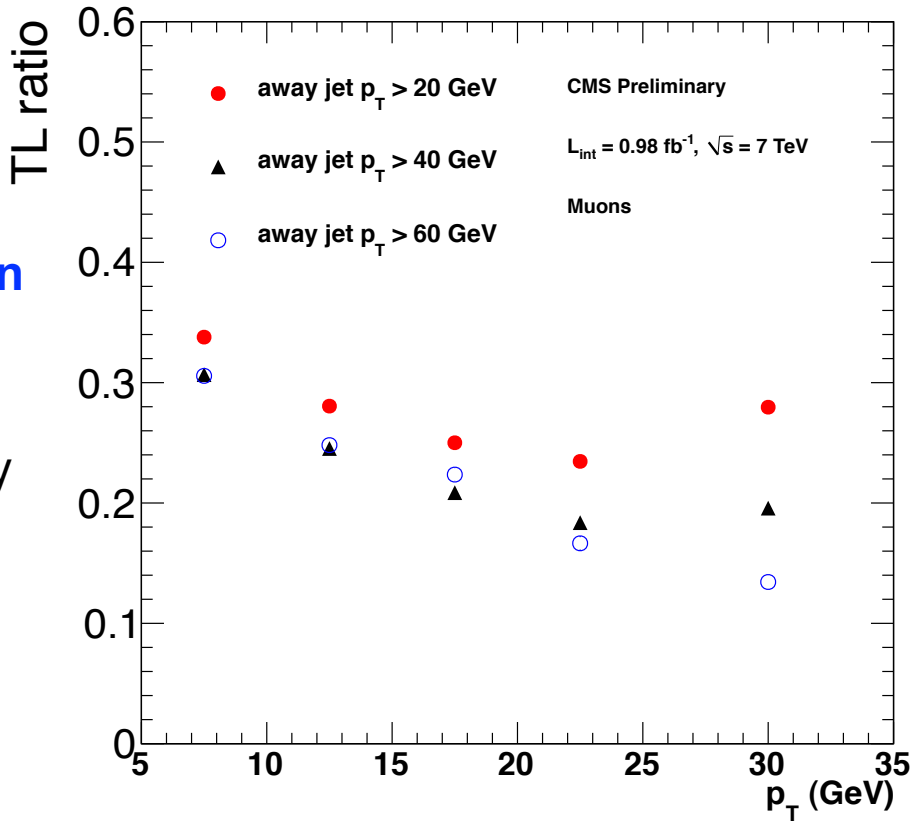
Background from Fake Leptons

- **We use data-driven methods based upon an extrapolation from the analysis selection to a region defined by looser lepton requirements.**

- **Two distinct methods:**
  - **Tight-Loose (TL) ratio Method (e, μ, τ_h) → A1, A2 and taus.**
    - Extrapolation of lepton selection in isolation.
    - Determine ratio from a QCD-enriched sample.
    - Estimate background with one and two fake leptons.
  - **Method B → different methods for one and two fake leptons.**
    - **B tag-and-probe method for background with one fake lepton.**
      - Extract lepton isolation templates.
      - Determine templates using a sample enriched in heavy flavor.
    - **Factorization Method for background with two fake leptons.**
      - Factorization of isolation and MET requirements.
      - Templates extracted using signal sample with relaxed isolation, MET.

Tight-Loose Ratio

TL ratio: $\epsilon(p_T, \eta) = \frac{\text{leptons passing tight selection}}{\text{leptons passing loose selection}}$

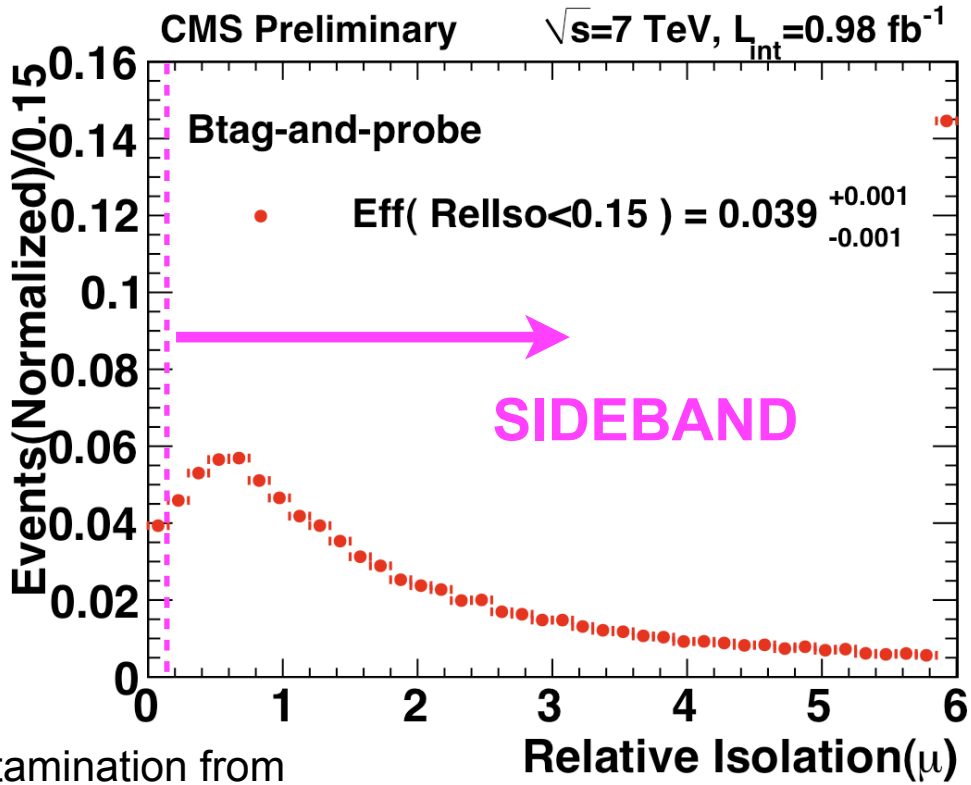


- **Measure TL ratio using events with a lepton passing the loose selection and an away side jet, to enrich the sample in fake leptons → leptons are jets.**
- The p_T of the away side jet acts as a proxy for the p_T of the parton from which the lepton candidate originated.
- **The TL ratio is dependent upon the away jet p_T requirement and the composition of the sample.**
- Estimate background from dilepton events with one (two) fake lepton(s) by counting events with one (two) lepton(s) passing the loose selection but failing the tight selection and one (zero) lepton(s) passing the tight selection and re-weighting by the appropriate TL ratio and summing.

B Tag-and-Probe Method

- Relax isolation on the least isolated lepton (probe).
- **Measure efficiency with tag-and-probe.**
  - Tag is b-tagged away jet → enrich sample in fake leptons.
  - Re-weight efficiency measured in data QCD events by the p_T, N_{jets} distributions expected from ttbar MC.

$$\epsilon = \frac{\text{leptons with relIso} < 0.15}{\text{leptons with relIso} > 0.15}$$



- **Isolation templates are used to estimate the background from events with one fake lepton:**

$$N^{SS}_{p-f} = (N_{sideband} - N^{SS}_{f-f})\varepsilon$$

Events passing tight selection with isolation relaxed on least isolated lepton.

Remove contamination from events with two fake leptons.

Factorization Method

- Apply a preselection by fully relaxing the MET and lepton isolation cuts in the region of interest. This provides a sample that is dominated by QCD.
- Measure the efficiencies by applying the cuts individually.

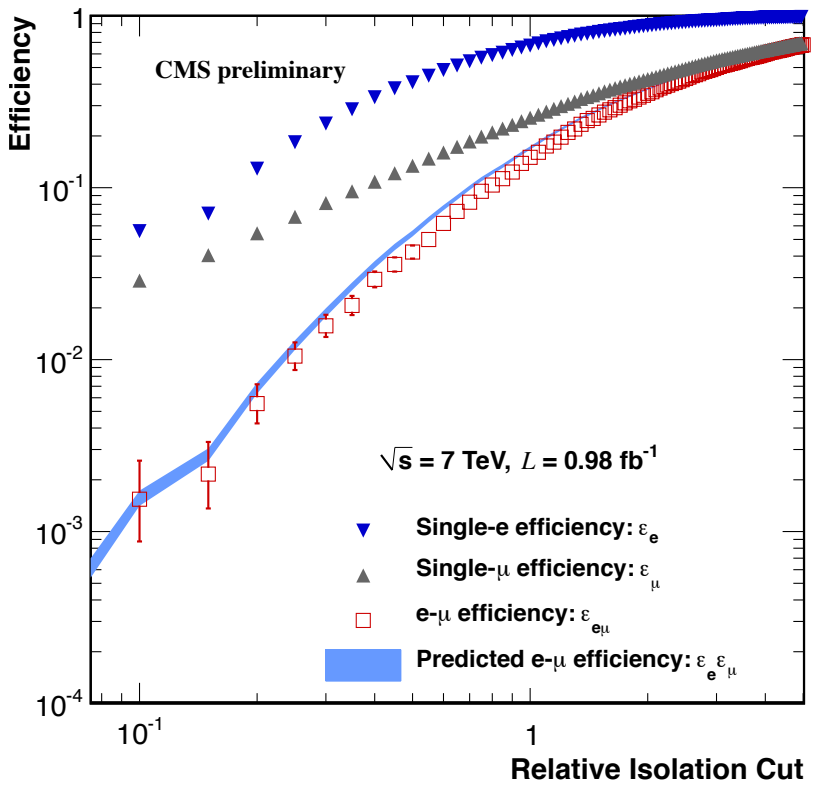
e.g. $$\epsilon_{\not{E}_T} = \frac{\text{events with no isolation requirements}}{\text{events with no isolation or } \not{E}_T \text{ requirements}}$$

- **Isolation, MET cuts independent → efficiencies factorize.**

$$N^{SS}_{f-f} = N^{\text{preselection}}_{\mu\mu} \cdot \epsilon_\mu \cdot \epsilon_\mu \cdot \epsilon_{\not{E}_T} + N^{\text{preselection}}_{ee} \cdot \epsilon_e \cdot \epsilon_e \cdot \epsilon_{\not{E}_T} + N^{\text{preselection}}_{e\mu} \cdot \epsilon_e \cdot \epsilon_\mu \cdot \epsilon_{\not{E}_T}$$

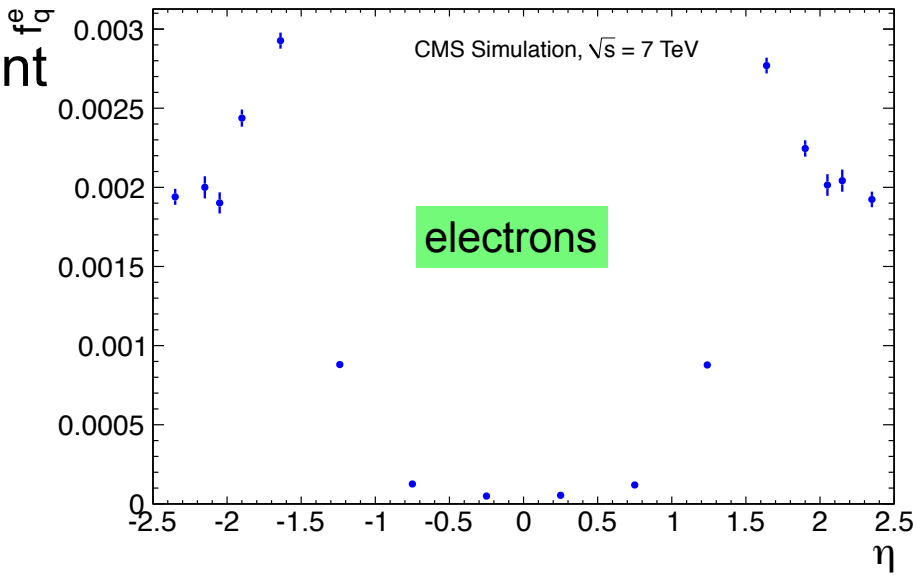
Background from events with two fake leptons.

Events from preselection sample with MET and lepton isolation requirements removed.

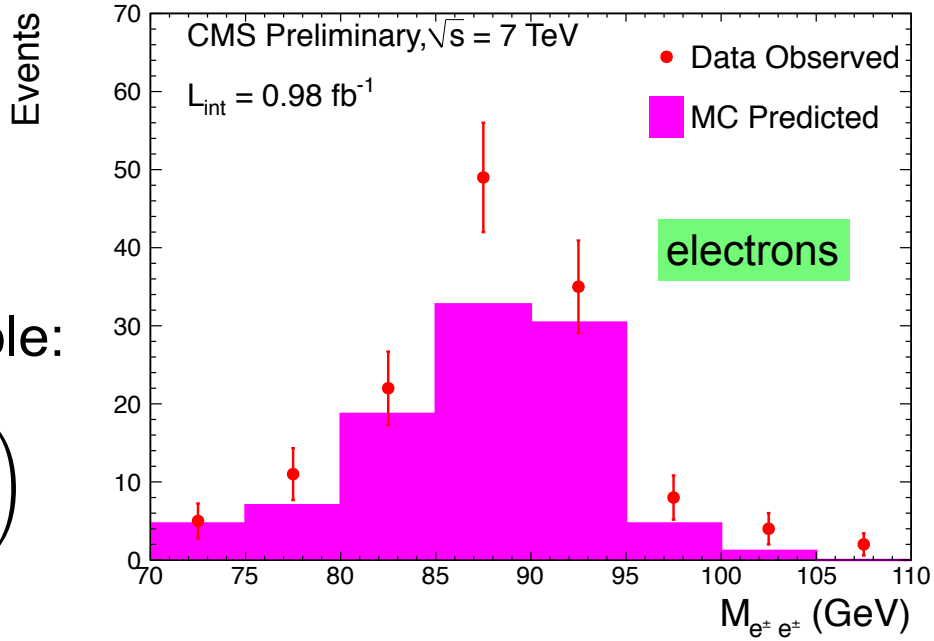


Background from Charge Mis-measurement

- Estimate the charge mis-measurement rate for electrons and taus using a same sign Z sample.
  - $\varepsilon(p_T, \eta) = N(SS) / \{N(SS) + N(OS)\}$
  - For electrons: $\varepsilon \approx 10^{-3} - 10^{-4}$
  - For taus: $\varepsilon \approx 0.07$
  - Rate for muons is negligible ($\approx 10^{-5}$)



- **Observe good agreement between same sign Z yield in data and that expected from simulation.**



- Estimate background from OS sample:

$$N_{\ell^\pm\ell^\pm} = N_{\ell_1^\pm\ell_2^\mp} \cdot \left(\frac{\epsilon_1}{1-\epsilon_1} + \frac{\epsilon_2}{1-\epsilon_2} \right)$$

Irreducible Background

- Remaining background from SM sources of prompt, isolated SS dileptons.
  - Use simulation to estimate the contribution from these processes.

Search Region (min. $H_T/\not{E}_T$)	WZ/ZZ	$2 \times (q\bar{q} \to W^{\pm})$	$t\bar{t}W$	$qq \to q'q'W^{\pm}W^{\pm}$	$V\gamma$	Charge misrec.	Total
Region 1 (400/120)							
Inclusive dileptons	0.06 ± 0.03	0.00 ± 0.002	0.28 ± 0.01	0.32 ± 0.02	0.00 ± 0.21	0.03 ± 0.01	0.7 ± 0.2
High–p_T dileptons	0.06 ± 0.03	0.00 ± 0.002	0.26 ± 0.01	0.31 ± 0.02	0.00 ± 0.21	0.03 ± 0.01	0.7 ± 0.2
τ dileptons	0.00 ± 0.01	0.00 ± 0.002	0.04 ± 0.02	0.05 ± 0.01	0.15 ± 0.15	0.8 ± 0.4	1.0 ± 0.4
Region 2 (400/50)							
Inclusive dileptons	0.10 ± 0.03	0.002 ± 0.002	0.62 ± 0.02	0.65 ± 0.03	0.00 ± 0.21	0.09 ± 0.02	1.5 ± 0.2
High–p_T dileptons	0.10 ± 0.03	0.002 ± 0.002	0.59 ± 0.02	0.63 ± 0.03	0.00 ± 0.21	0.09 ± 0.02	1.4 ± 0.2
Region 3 (200/120)							
Inclusive dileptons	0.20 ± 0.05	0.002 ± 0.002	0.86 ± 0.02	0.62 ± 0.02	0.00 ± 0.21	0.08 ± 0.01	1.8 ± 0.2
High–p_T dileptons	0.20 ± 0.05	0.00 ± 0.002	0.79 ± 0.02	0.57 ± 0.02	0.00 ± 0.21	0.08 ± 0.01	1.6 ± 0.2
Region 4 (80/100)							
High–p_T dileptons	0.43 ± 0.07	0.004 ± 0.003	1.51 ± 0.02	0.93 ± 0.03	0.00 ± 0.21	0.28 ± 0.03	3.2 ± 0.2

$t\bar{t}W,\ qq \to q'q'W^{\pm}W^{\pm}$

WZ/ZZ measured at CMS, Tevatron.
→ small background
→ well understood

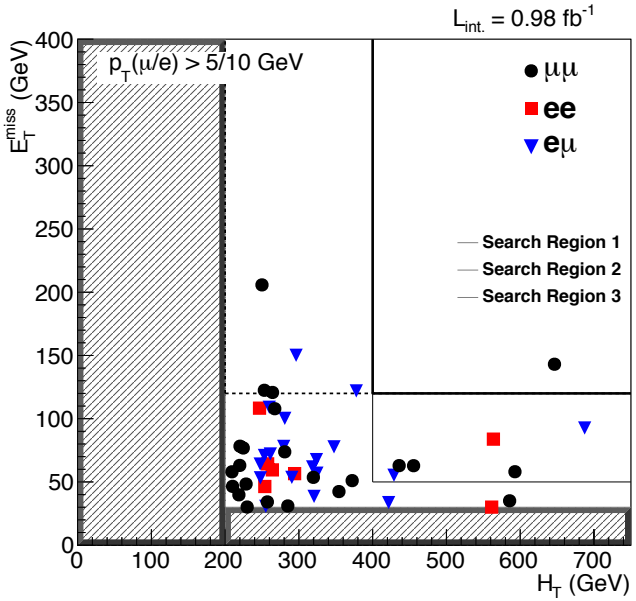
Double parton scattering background is small.

Rare SM processes, never measured before → opportunity for new observation!

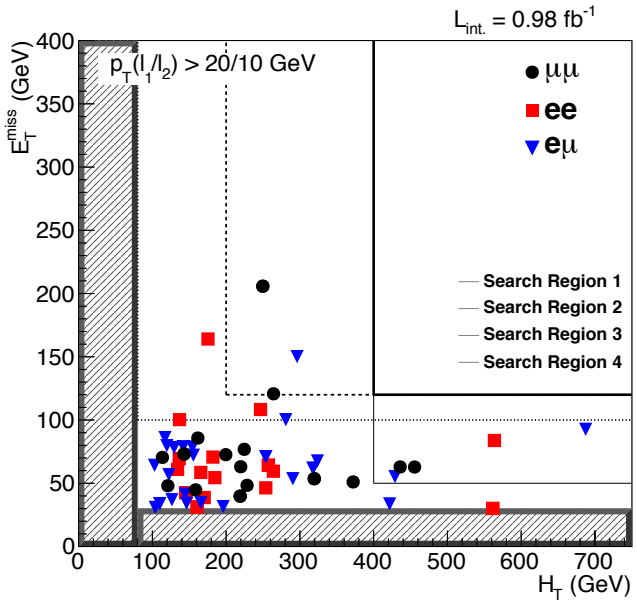
Contribute significantly to high MET, high H_T signal regions → understanding or suppressing these backgrounds may become important for same sign searches with more data.

Yields in H_T-MET Plane

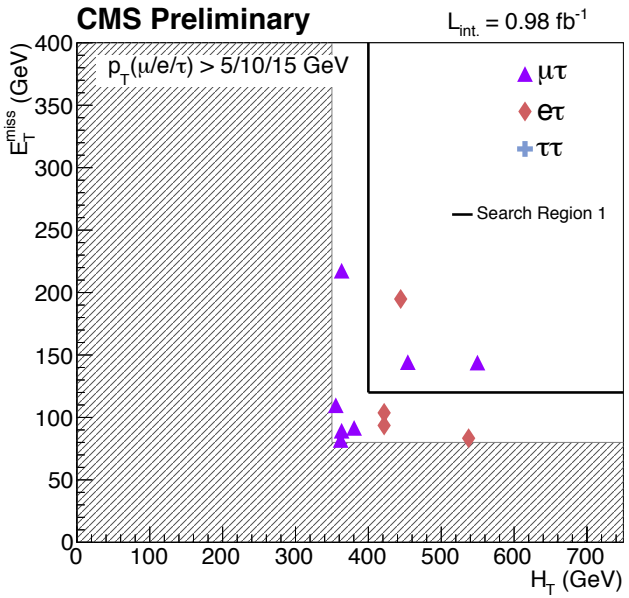
inclusive-p_T



high-p_T

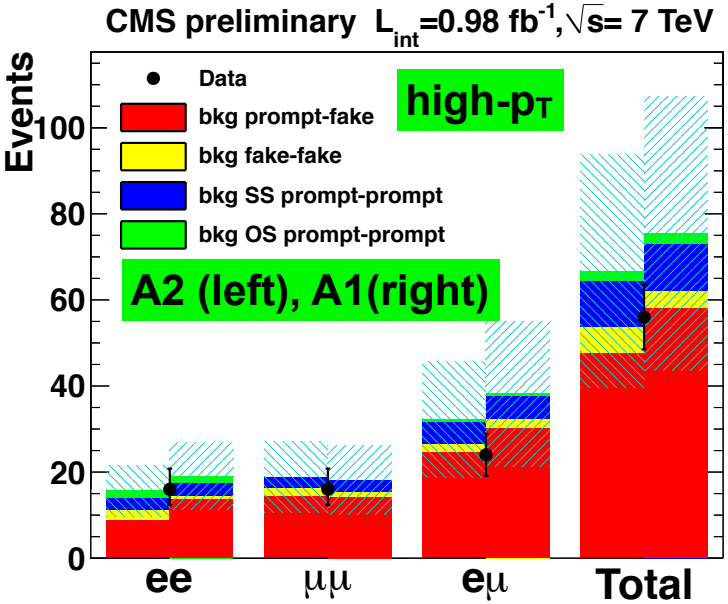


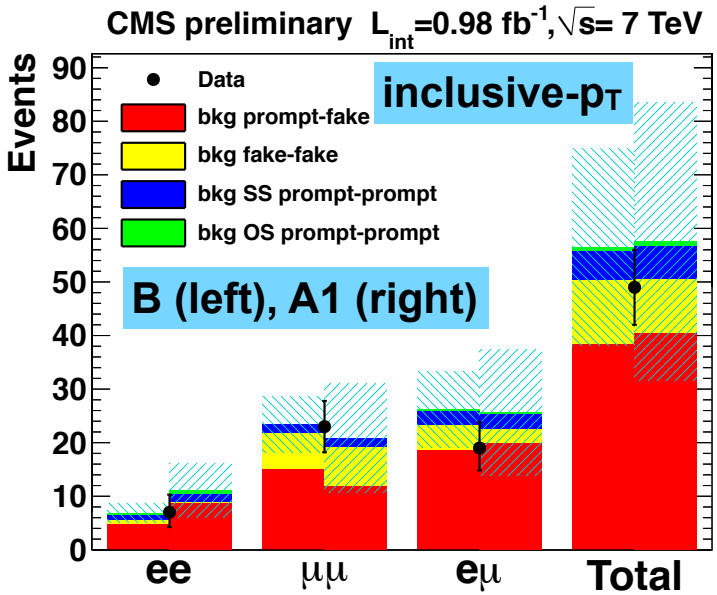
taus



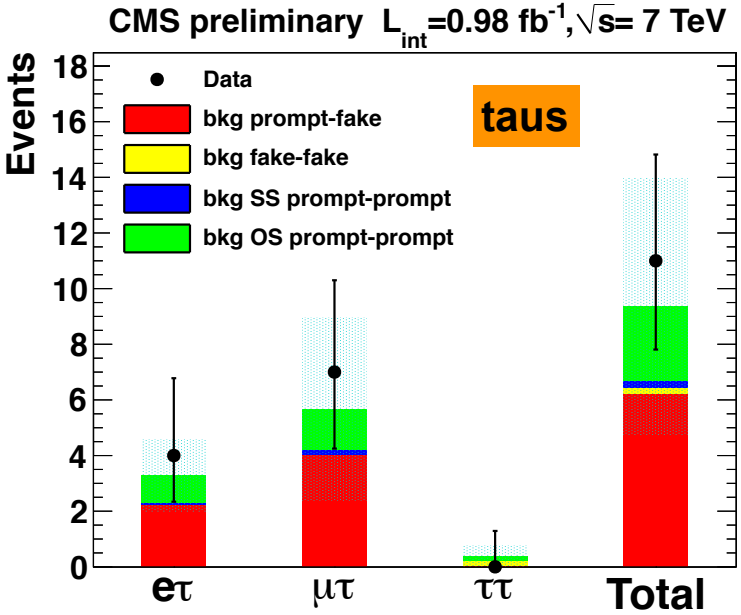
- H_T vs. MET scatter plots for the three control regions in data. **For all three searches, the events in data cluster at low H_T and low MET, a region expected to be dominated by background.**

Control Regions - Results



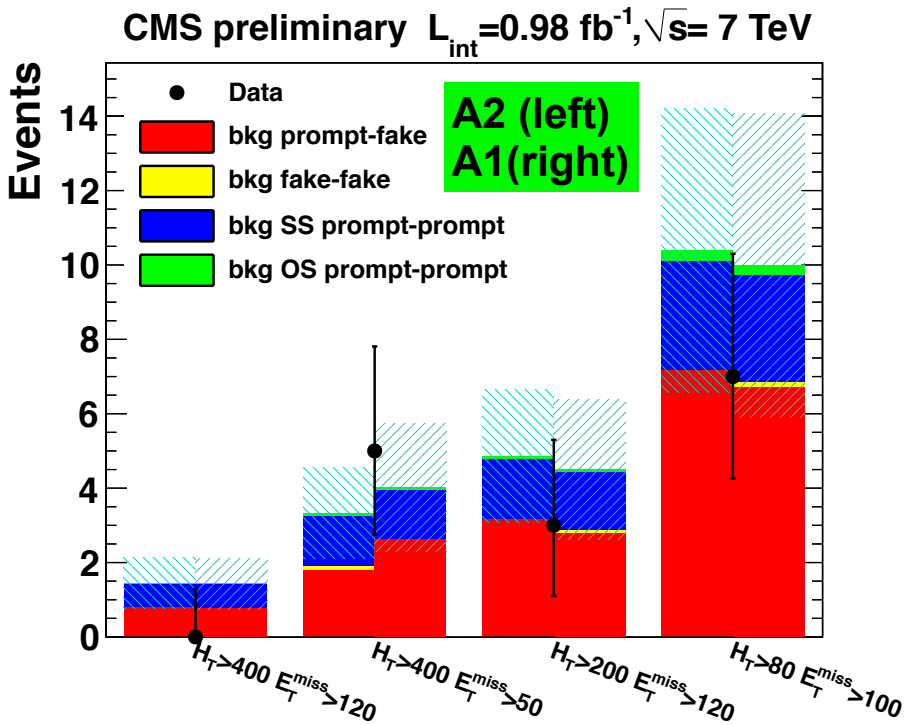


- Breakdown of background estimates in the three baseline regions. **For inclusive- and high-p_T dileptons, observe good agreement for all components of the background estimate → this gives confidence that our background estimates work. Let's go look for new physics.**

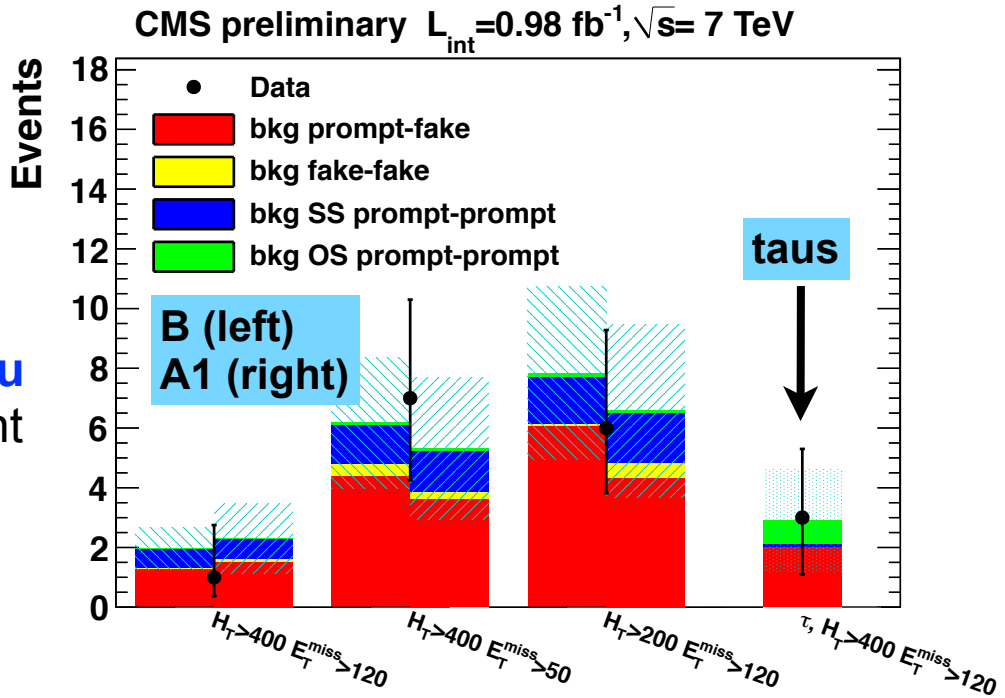


Results

We find no evidence for new physics.



Left, top: Yields and background estimates for the **high-p_T search regions**. We find good agreement between the yields and the background estimates. The two different background estimates agree well in all regions.



Right, bottom: Yields and background estimates for the **inclusive-p_T and the tau search regions**. We find good agreement between the yields and the background estimates. For the inclusive-p_T searches, the two different background estimates agree well in all regions.

Yields and Limits: High-p_T Searches

Search Region (minimum $H_T/\not{E}_T$)	ee	$\mu\mu$	$e\mu$	Total	95% CL UL yield
Region 1 (400/120)					
Predicted background by A1	0.4 ± 0.3	0.4 ± 0.3	0.7 ± 0.4	1.4 ± 0.7	
Predicted background by A2	0.7 ± 0.5	0.4 ± 0.3	0.4 ± 0.3	1.4 ± 0.7	
Observed	0	0	0	0	3.0
Region 2 (400/50)					
Predicted background by A1	1.4 ± 0.8	1.3 ± 0.8	1.3 ± 0.6	4.0 ± 1.7	
Predicted background by A2	1.5 ± 0.8	0.8 ± 0.4	1.0 ± 0.5	3.3 ± 1.2	
Observed	1	2	2	5	7.5
Region 3 (200/120)					
Predicted background by A1	1.2 ± 0.7	1.5 ± 0.8	1.8 ± 0.8	4.5 ± 1.9	
Predicted background by A2	1.3 ± 0.7	1.8 ± 0.8	1.8 ± 0.7	4.9 ± 1.8	
Observed	0	2	1	3	5.2
Region 4 (80/100)					
Predicted background by A1	2.5 ± 1.2	2.6 ± 1.2	4.9 ± 2.2	10 ± 4	
Predicted background by A2	2.4 ± 1.0	3.6 ± 1.6	4.4 ± 1.6	10 ± 4	
Observed	3	2	2	7	6.0

- **Good agreement between observed yields and estimated backgrounds.** Set model independent limits using the background estimate that provides the more conservative limit.

Signal Efficiency Model

- Parameterization of the efficiencies of the lepton, MET and HT selections.
- Model developed using generator level SUSY sample (LM6).
- How to use:
  - Implement model of interest in favorite Monte Carlo.
  - Apply analysis selections at generator-level.
  - Use efficiency model to scale generator-level yields to “data-level”.
  - Compare expected yield from scaled MC with 95% CL UL yield.

- Isolation corrections (from SUS-10-004):

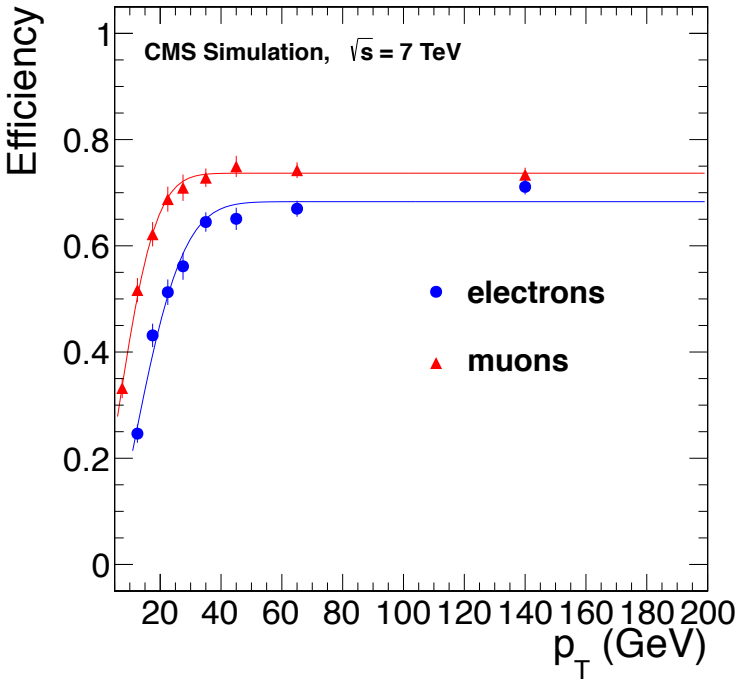
$$\Delta\epsilon = -0.10\frac{<n>-25}{15},$$

where $<n>$ is the average number of stable charged particles with $p_T > 3$ GeV and $|\eta| < 2.4$

How to use:

Compare average number of stable charged particles at generator level for the model of interest with LM6.

For every 15 additional (fewer) charged particles, decrease (increase) the lepton efficiency by 10%.

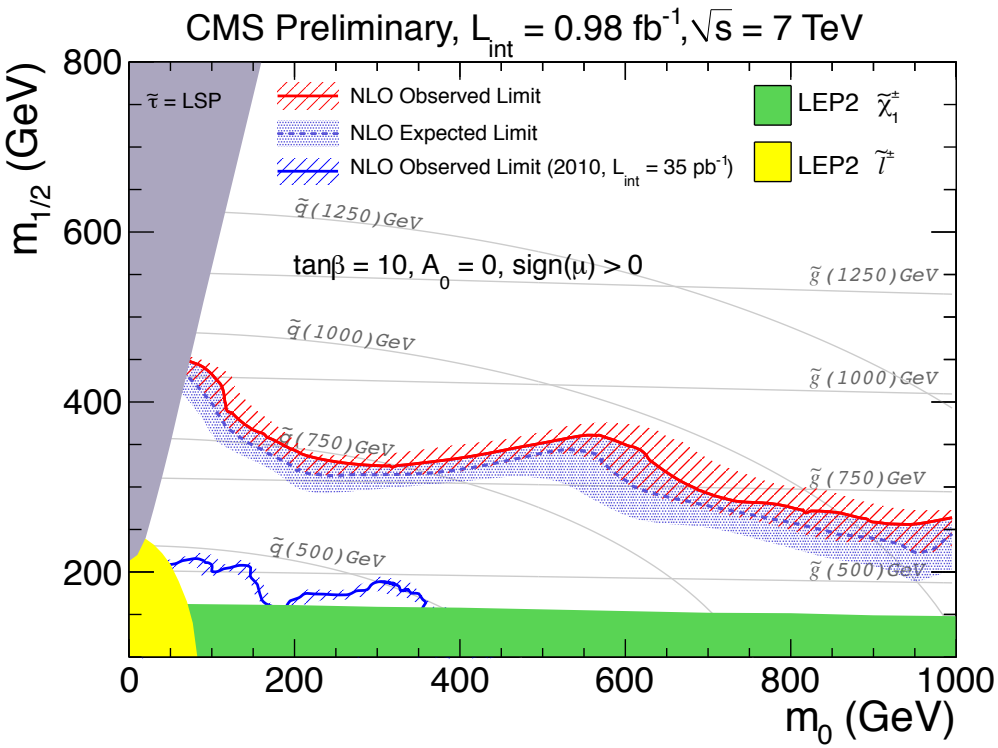


Exclusion Curve

- Interpret the upper limits in the context of a particular model.
  - Here, consider high-p_T search with H_T > 400 GeV, MET > 120 GeV.

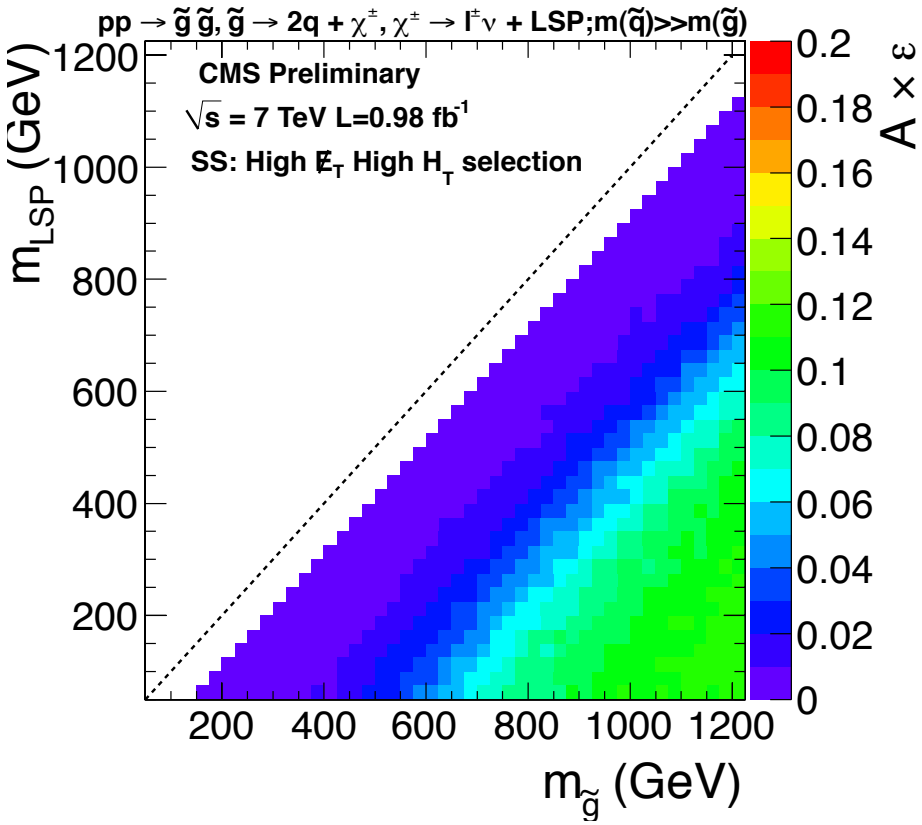
- **CMSSM** is a benchmark SUSY model used by the LHC and the Tevatron.

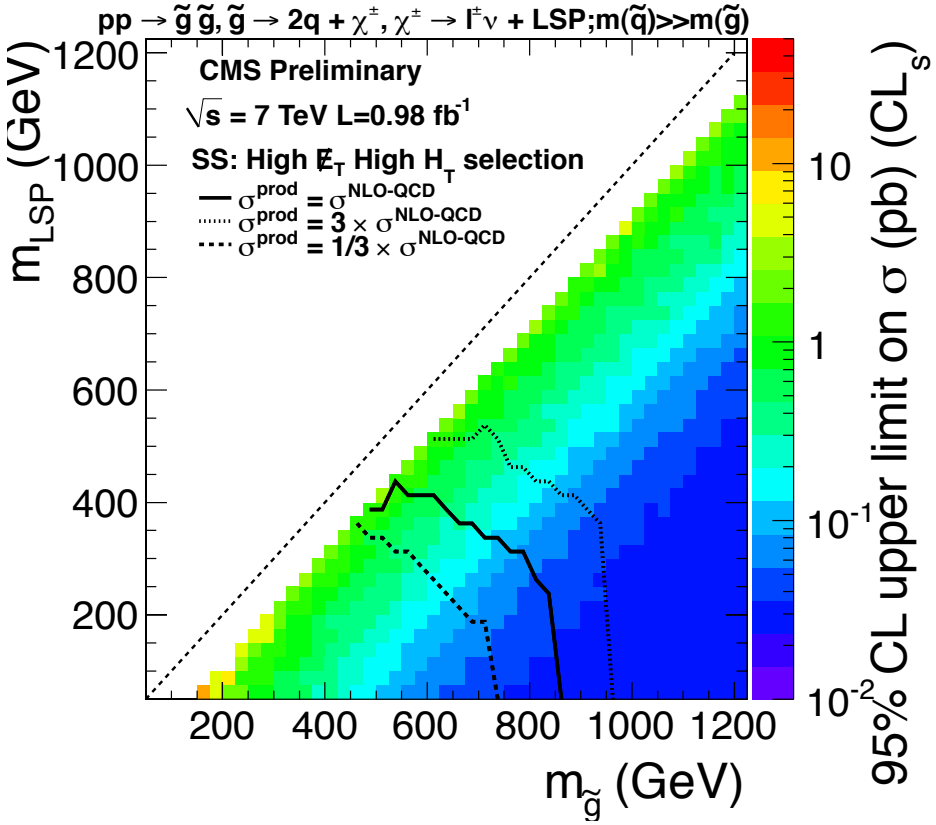
- An exclusion is made in the **m_0-$m_{1/2}$** plane (2 of the 5 parameters in the CMSSM Lagrangian).



- **The Summer 2011 CMS exclusion improves on previous results from CMS and the Tevatron experiments.**

Simplified Models





- Small SM backgrounds allow same-sign searches to use lower H_T and MET requirements (and low-p_T leptons) which enables these searches to probe models with compressed spectra (see talk by Mariarosaria D'alfonso on Tuesday for more details).

Summary

- We searched for new physics in the same sign dilepton + jets + MET final state in *pp* collisions at 7 TeV using a dataset corresponding to 0.98 fb^{-1}.
  - **We observe no evidence for new physics. We set model-independent limits on new physics in the signal regions.**
  - **We validate the data-driven background estimates using a control region and find good agreement. Multiple, independently executed background estimates provide confirmation in each region.**
  - We provide a parameterization of the signal acceptance to allow for the interpretation of the observed experimental limits in the context of a broad range of models.

Backup Slides

Control Regions - Results

- Cross check background estimates by applying them to a control region that is enriched in background.

p_T > 20, 10 GeV
H_T > 80 GeV, MET > 30 GeV

p_T > 5(10) GeV for μ(e)
H_T > 200 GeV, MET > 30 GeV

p_T > 5(10)[15] GeV for μ(e)[τau]
H_T > 350 GeV, MET > 80 GeV

Baseline Region	ee	$\mu\mu$	$e\mu$	total
inclusive dileptons				
predicted BG (B)	7 ± 3	23 ± 5	26 ± 7	56 ± 18
predicted BG (A1)	11 ± 5	21 ± 10	26 ± 12	58 ± 26
observed	7	23	19	49
high–p_T dileptons				
predicted BG (A1)	19 ± 8	18 ± 8	38 ± 17	75 ± 32
predicted BG (A2)	16 ± 6	19 ± 8	32 ± 13	67 ± 27
observed	16	16	24	56

Baseline Region	$e\tau_h$	$\mu\tau_h$	$\tau_h\tau_h$	total
τ_h dileptons				
predicted BG	3.2 ± 1.3	5.6 ± 3.3	0.3 ± 0.4	9.1 ± 3.6
observed	4	7	0	11

- **Good agreement between observation and prediction. For inclusive- and high-p_T dileptons, find good agreement between the different background estimates.**

Yields and Limits: Inclusive Searches

Search region (minimum $H_T/\not{E}_T$)	ee	$\mu\mu$	$e\mu$	Total	95% CL UL yield
Region 1 (400/120)					
Predicted background by B	0.2 ± 0.1	0.9 ± 0.3	0.9 ± 0.3	2.0 ± 0.7	
Predicted background by A1	0.4 ± 0.4	1.2 ± 0.8	0.7 ± 0.4	2.3 ± 1.2	
Observed	0	1	0	1	3.7
Region 2 (400/50)					
Predicted background by B	1.0 ± 0.4	2.3 ± 0.7	3.0 ± 1.0	6.2 ± 2.2	
Predicted background by A1	1.3 ± 0.7	2.5 ± 1.5	1.4 ± 0.7	5.3 ± 2.4	
Observed	1	4	2	7	8.9
Region 3 (200/120)					
Predicted background by B	0.8 ± 0.4	3.6 ± 1.3	3.4 ± 1.3	7.8 ± 2.9	
Predicted background by A1	1.5 ± 0.9	3.0 ± 1.6	2.1 ± 1.0	6.6 ± 2.9	
Observed	0	4	2	6	7.3

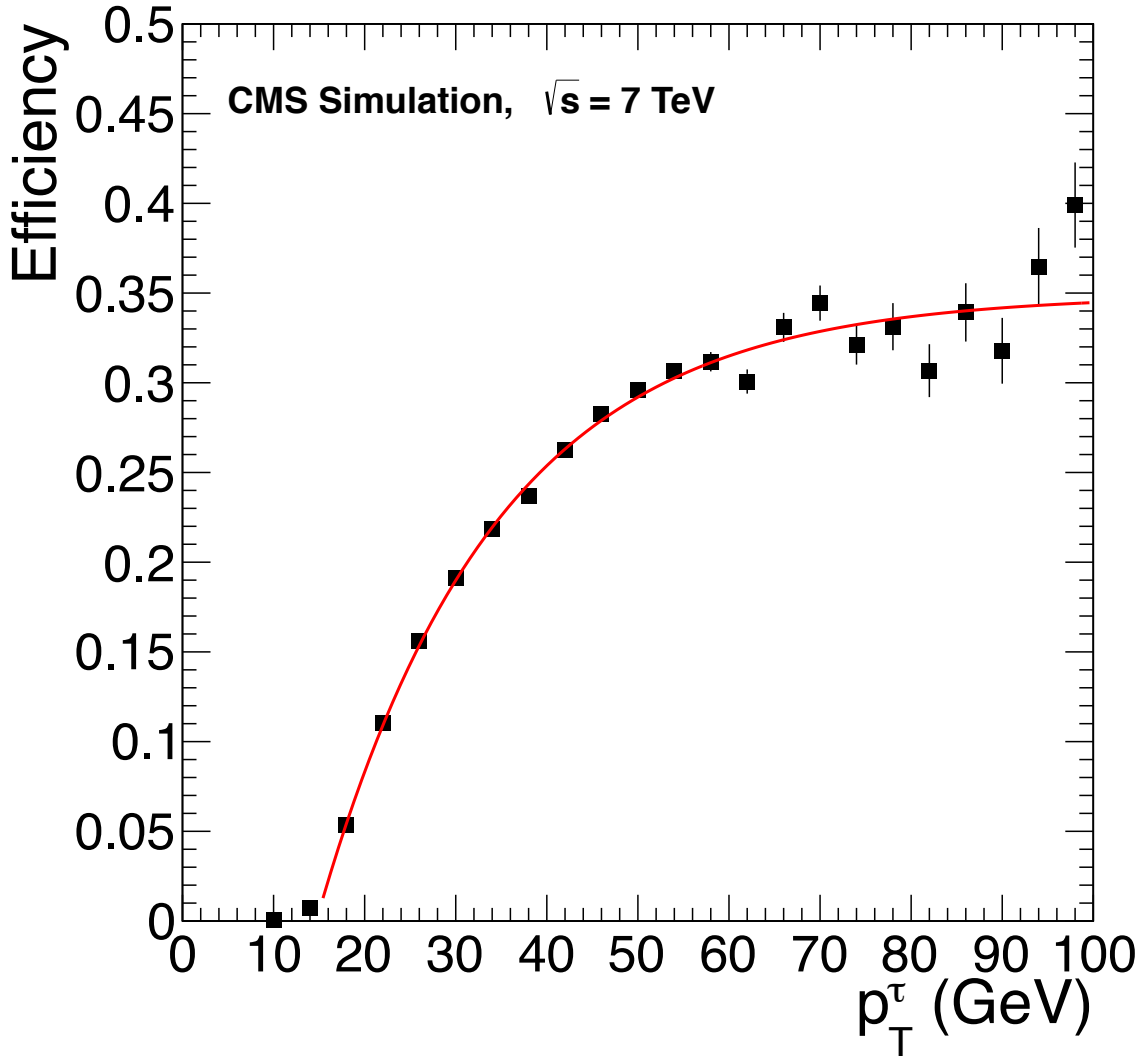
- **Good agreement between observed yields and estimated backgrounds.** Set model independent limits using the background estimate that provides the more conservative limit.

Yields and Limits: Tau Searches

Search Region (minimum $H_T/\not{E}_T$)	$e\tau$	$\mu\tau$	$\tau\tau$	Total	95% CL UL yield
Region 1 (400/120)					
Predicted background	1.1 ± 0.4	1.8 ± 1.4	0.0 ± 0.2	2.9 ± 1.7	
Observed	1	2	0	3	5.8

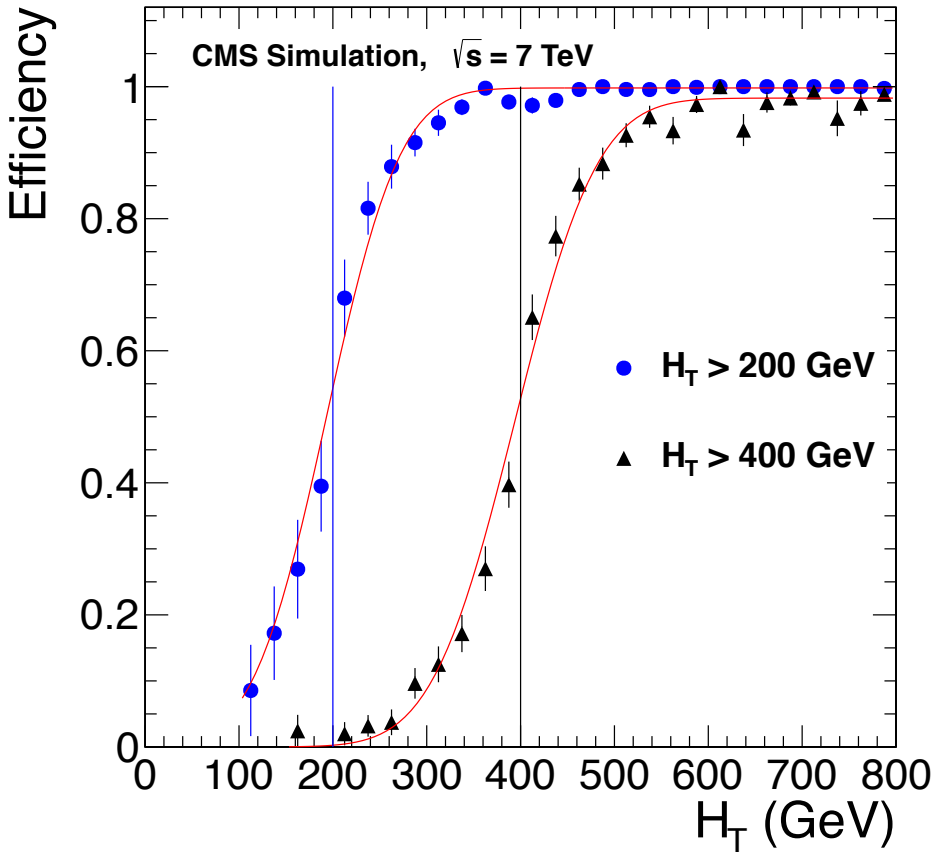
- **Good agreement between observed yields and estimated backgrounds.**

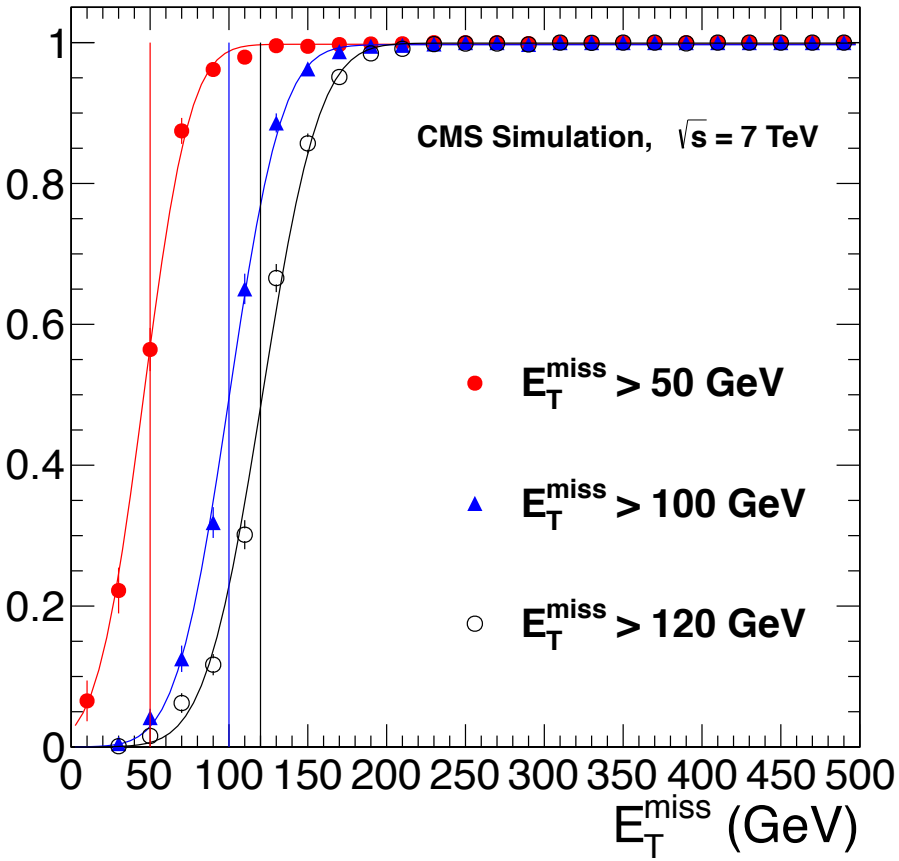
Signal Efficiency Model - Taus



- Hadronic tau selection efficiency.

Signal Efficiency Model - H_T and MET





- Efficiencies of the H_T (left) and MET (right) selections.