Search for Top Partners using same-sign dilepton events

A. Avetisyan, T. Bose, S. Lewin (Boston University)
S. Bhattacharya, M. Narain (Brown University)
P. Azzi, P. Checchia, M. Nespolo (INFN - Sezione di Padova)
M. Abis, S. Vanini (Universita' di Padova & INFN)

Pre-Approval

June 1st, 2012

Top Partners

- Common prediction of different theories
  - Couple to 3rd generation quarks
  - Solve hierarchy problem

- Can be found in
  - Higgs-less models
  - Minimal Composite Higgs
  - Extra dimensions (KK gluons)

- Theoretical descriptions
  - ***Contino & Servant***, JHEP 0806:026 (2008)
  - ***Mrazek & Wulzer***, Phys. Rev. D 81, 075006 (2010)

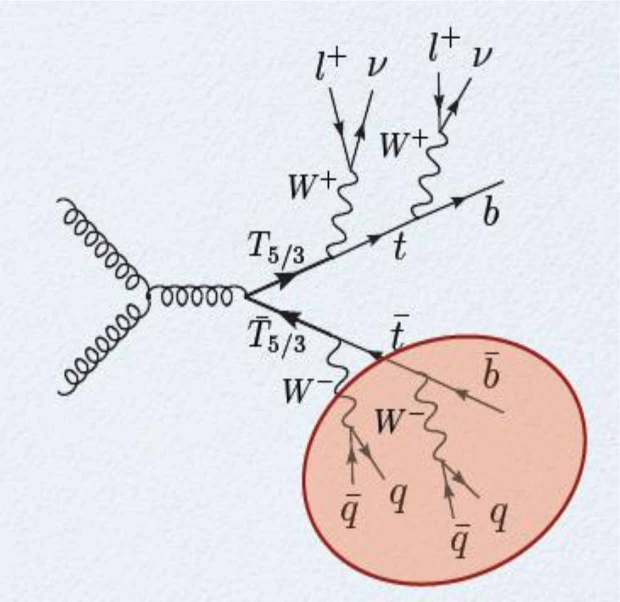
- Previous CMS PAS (MC-based): EXO-08-008

Top Partners Model

- Our model of top partners:
  - $T_{5/3}$ with Q_e = 5/3 and B with Q_e = -1/3 decay into W and top

- Focus on pair production of $T_{5/3}$
  - In Mrazek & Wulzer model, B is typically significantly more massive

- $T_{5/3}$ signature: $\boxed{l^{\pm}l^{\pm} + n \text{ jets}}$



Signal Samples

- Initial samples suffered from MadGraph bug
  - Worked with theorists and MG experts to understand this
- $T_{5/3}$ masses from 400 to 750 GeV in intervals of 50 GeV
- Generated using MadGraph
  - Two same-signed W's forced to decay to leptons (e, μ, τ)
  - The other two W's decay inclusively
- Fall11 MC production
  - Using CMSSW_4_2_X
  - Initially used private FastSim, but everything is now with centrally produced FullSim
- TopLikeBSM pattuples (v9) used for both signal and background

Signal Samples

$T_{5/3}$ Mass (GeV)	σ (pb)	$\sigma\times$ BR (pb)	Number of events
400	1.405	0.295	86,205
450	0.662	0.139	86,211
500	0.330	0.069	86,684
550	0.171	0.036	86,724
600	0.092	0.019	86,965
650	0.051	0.011	87,592
700	0.029	0.006	88,145
PAS 750	0.017	0.004	88,410

- NNLO cross-section calculated with HATHOR
- Branching ratio is 0.21

Background Samples

Process PAS	MC Generator	σ (pb)	Number of events
Background Samples:			
$t\bar{t}$	MADGRAPH	164.4 (NNLO)	3,701,947
$W(\rightarrow)\ell\nu$+jets	MADGRAPH	31314 (NNLO)	77,105,816
Drell Yan - $Z/\gamma^*(\rightarrow \ell\ell)$+jets	MADGRAPH	3048 (NNLO)	36,277,961
WW	MADGRAPH	4.882 (NLO)	1,197,558
WZ	MADGRAPH	0.879 (NLO)	1,221,134
ZZ	MADGRAPH	0.076 (NLO)	1,185,188
Single Top (tW)	POWHEG	15.74 (NNLO)	1,624,374
Single Top ($t - channel$)	POWHEG	64.57 (NNLO)	5,844,997
Single Top ($s - channel$)	POWHEG	4.63 (NNLO)	397,951
$W\gamma \rightarrow e\nu\gamma$	MADGRAPH	137.3 (NLO)	524,503
$W\gamma \rightarrow \mu\nu\gamma$	MADGRAPH	137.3 (NLO)	521,774
W^+W^+	MADGRAPH	0.165	130,000
W^-W^-	MADGRAPH	0.055	160,000
WWW	MADGRAPH	0.038	1,201,777
$t\bar{t}W$	MADGRAPH	0.169 (NLO)	1,029,608
$t\bar{t}Z$	MADGRAPH	0.139	793,155

Summer11 except for ttW/Z, WWW and same-sign WW (the latter are same as for SUS-010-11)

Datasets

Dataset	Run range
/DoubleMuon/Run2011A-May10ReReco-v1/AOD	160329-163869
/DoubleMuon/Run2011A-PromptReco-v4/AOD	165071-168437
/DoubleMuon/Run2011A-05AugReReco-v1/AOD	170053-172619
/DoubleMuon/Run2011A-PromptReco-v6/AOD	172620-175770
/DoubleMuon/Run2011B-PromptReco-v1/AOD	175832-180296
/DoubleElectron/Run2011A-May10ReReco-v1/AOD	160329-163869
/DoubleElectron/Run2011A-PromptReco-v4/AOD	165071-168437
/DoubleElectron/Run2011A-05AugReReco-v1/AOD	170053-172619
/DoubleElectron/Run2011A-PromptReco-v6/AOD	172620-175770
/DoubleElectron/Run2011B-PromptReco-v1/AOD	175832-180296
/MuEG/Run2011A-May10ReReco-v1/AOD	160329-163869
/MuEG/Run2011A-PromptReco-v4/AOD	165071-168437
/MuEG/Run2011A-05AugReReco-v1/AOD	170053-172619
/MuEG/Run2011A-PromptReco-v6/AOD	172620-175770
/MuEG/Run2011B-PromptReco-v1/AOD	175832-180296

AN

Table 1: Data samples used for the analysis

Golden JSON, 5.0 fb^{-1}

Triggers

HLT_DoubleMu7_v1,2	160404-163869
HLT_Mu13_Mu8_v2,3,4,6,7	165088-178380
HLT_Mu17_Mu8_v10,11	178420-180252
HLT_Ele17_CaloIdL_CaloIsoVL_Ele8_CaloIdL_CaloIsoVL_v1,2,3,4,5,6	160404-167913
HLT_Ele17_CCTT_Ele8_CCTT_v6,7,8,9,10	170249-180252
HLT_Mu10_Ele10_CaloIdVL_v2,3,4 or	160404-163869
HLT_Mu17_Ele8_CaloIdVL_v1,2,3,4,5,6,8 or	165088-173198
HLT_Mu17_Ele8_CaloIdT_CaloIsoVL__v4,7,8 or	173236-180252
HLT_Mu8_Ele17_CaloIdL_v1,2,3,4,5,6 or	160404-167913
HLT_Mu8_Ele17_CaloIdT_CaloIsoVL_v3,4,7,8	170249-180252

PAS

- CCTT = CaloIdT_CaloIsoVL_TrkIdVL_TrkIsoVL
- Scale factors are of order 1% (from AN-11-333)
  - We apply a systematic uncertainty of the same order

Event Selection

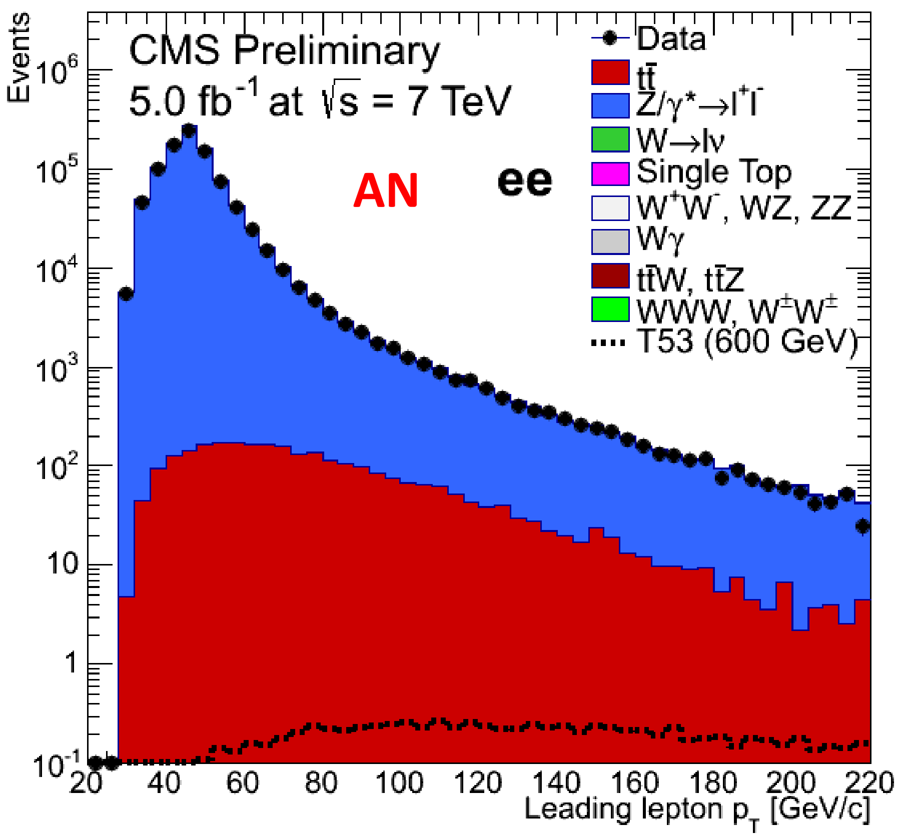
- Event cleanup and vertex reconstruction is done at the level of TopLikeBSM pattuples
- Scraping veto: reject event if fraction of high purity tracks is less than 25% in an event with at least 10 tracks
- Require at least one Primary Vertex
  - More than 4 degrees of freedom
  - Less than 24 cm away from nominal interaction point in z
  - Less than 2 cm away radially
- HBHE noise filtering

Object Reconstruction: Electrons

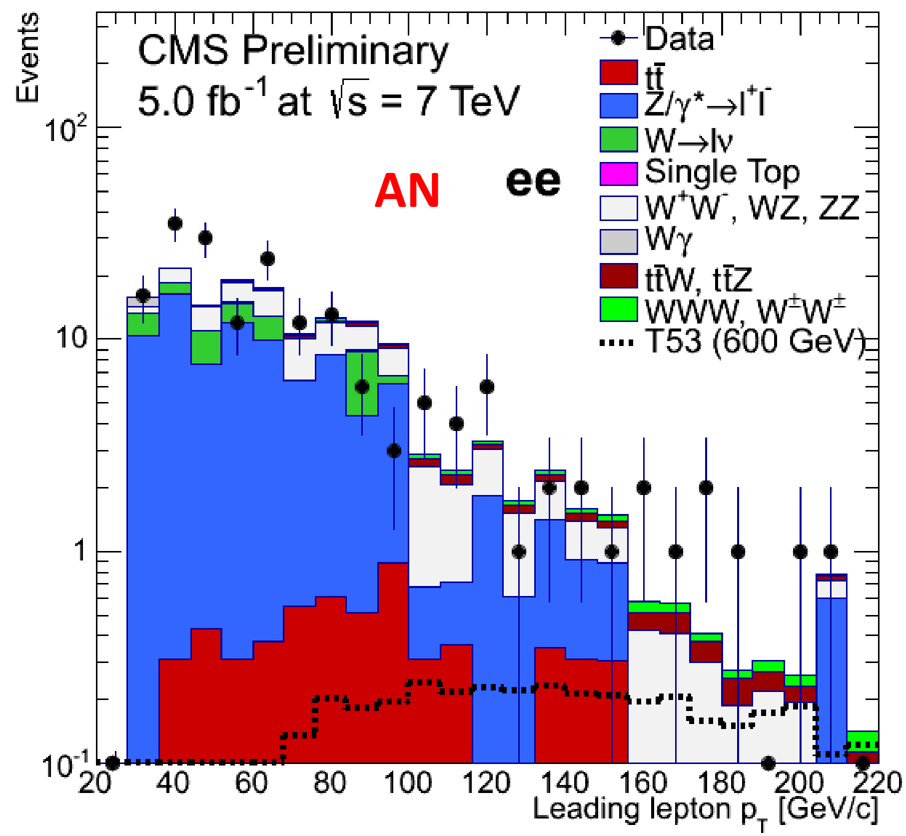
- Standard 2011 top quark electron selection except for charge consistency check (specific to same-sign analyses)

  - $p_T > 30$ GeV
  - $|\eta| < 2.4$ except EB-EE gap ($1.4442 < |\eta_{sc}| < 1.566$)
  - HyperTight1MC Cut in Categories electron ID
  - PF-based RelIso < 0.15
  - Conversion rejection
  - Transverse IP < 0.02 cm
  - Gsf-Ctf-ScPix charge consistency

Electron p_T



Two good electrons (no further requirements)



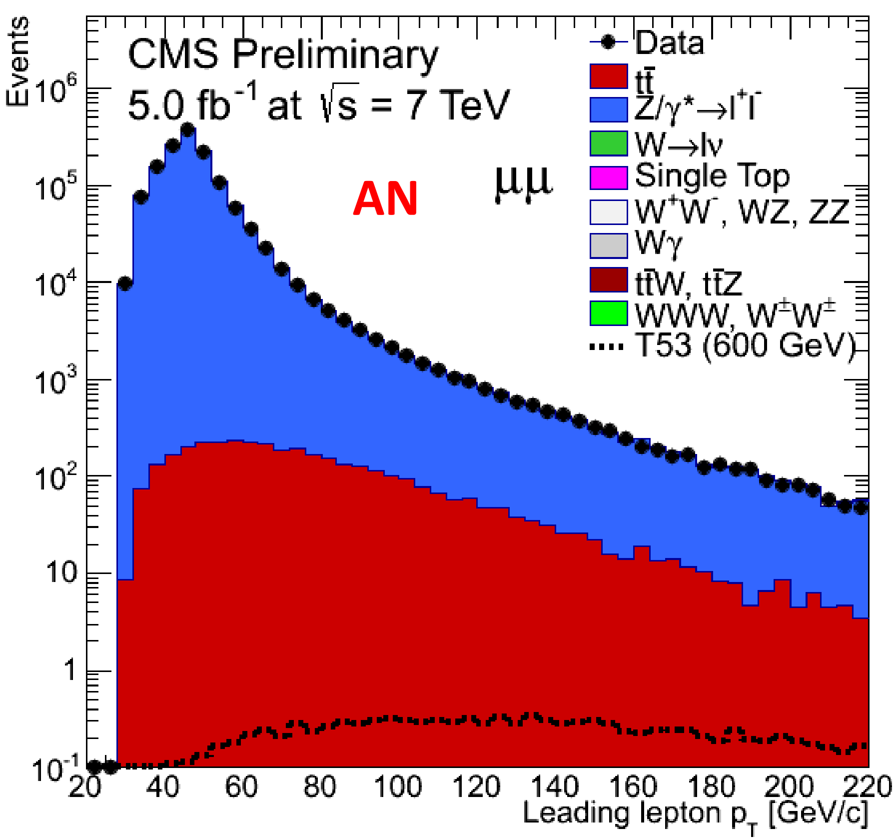
Two good electrons + same sign requirement + Z-veto

Object Reconstruction: Muons

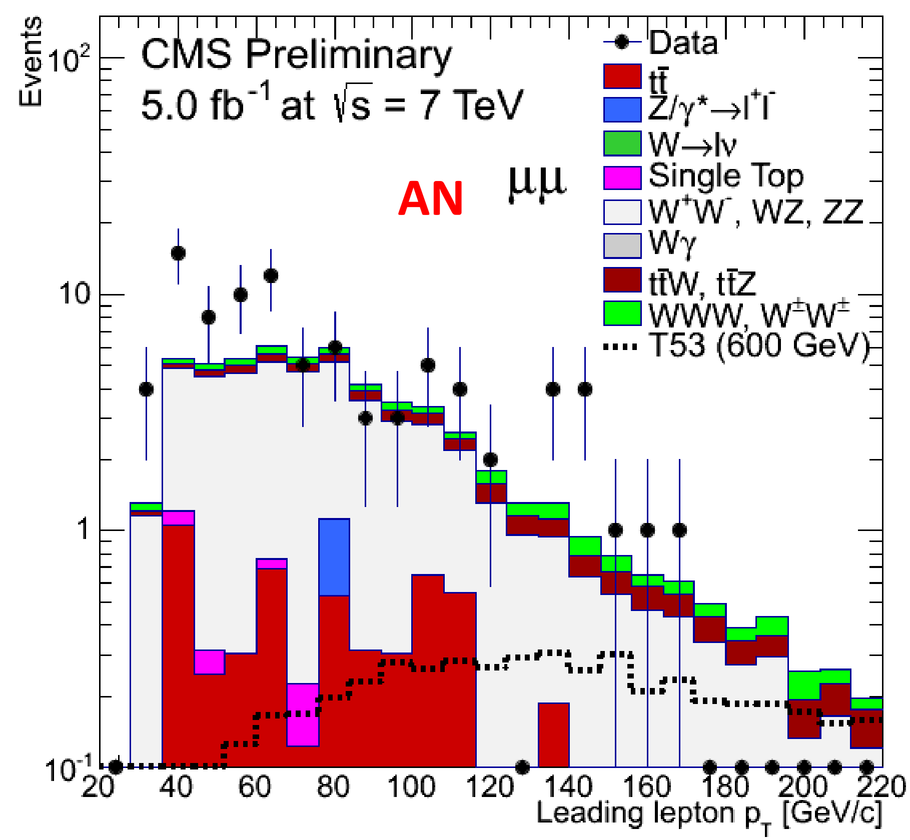
- Standard 2011 top quark muon selection

  - $p_T > 30$ GeV
  - $|\eta| < 2.4$
  - Global Muon and Tracker Muon
  - PF-based RelIso < 0.20
  - χ^2 / ndof < 10
  - Transverse IP < 0.02 cm
  - N Muon hits > 0
  - N Pixel hits > 0
  - N valid silicon hits > 10
  - N chambers with matched segments > 1

Muon p_T



Two good muons (no further requirements)



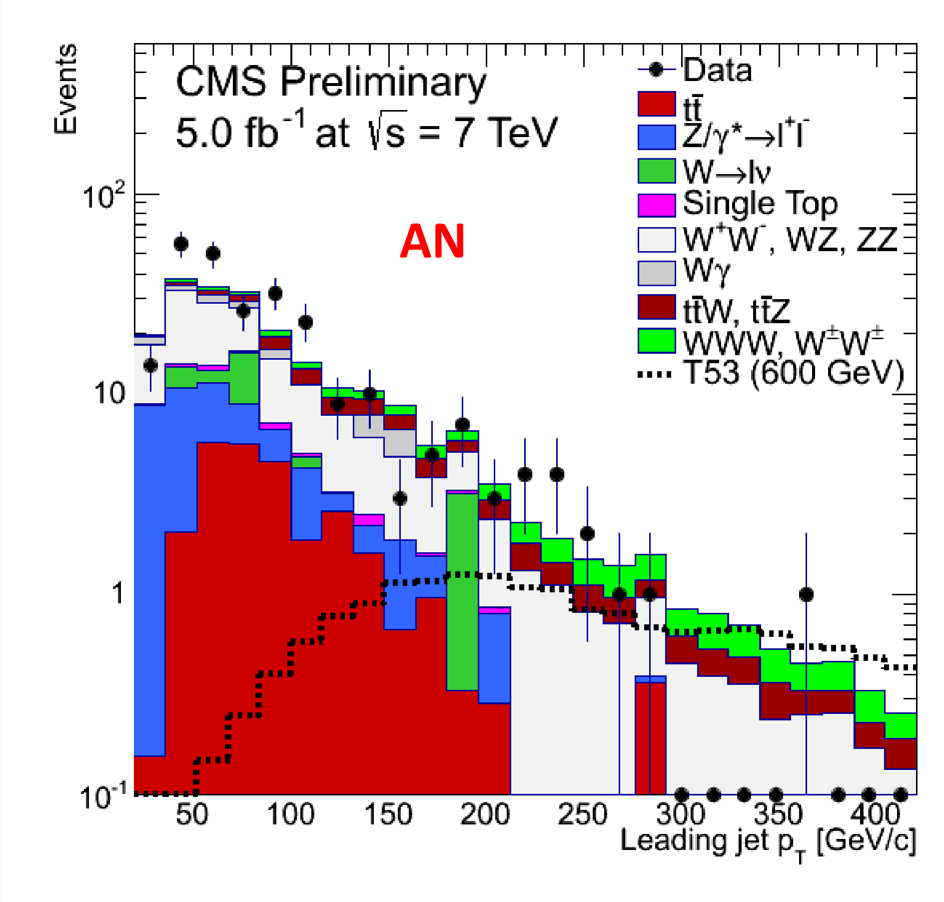
Two good muons + same sign requirement + Z-veto

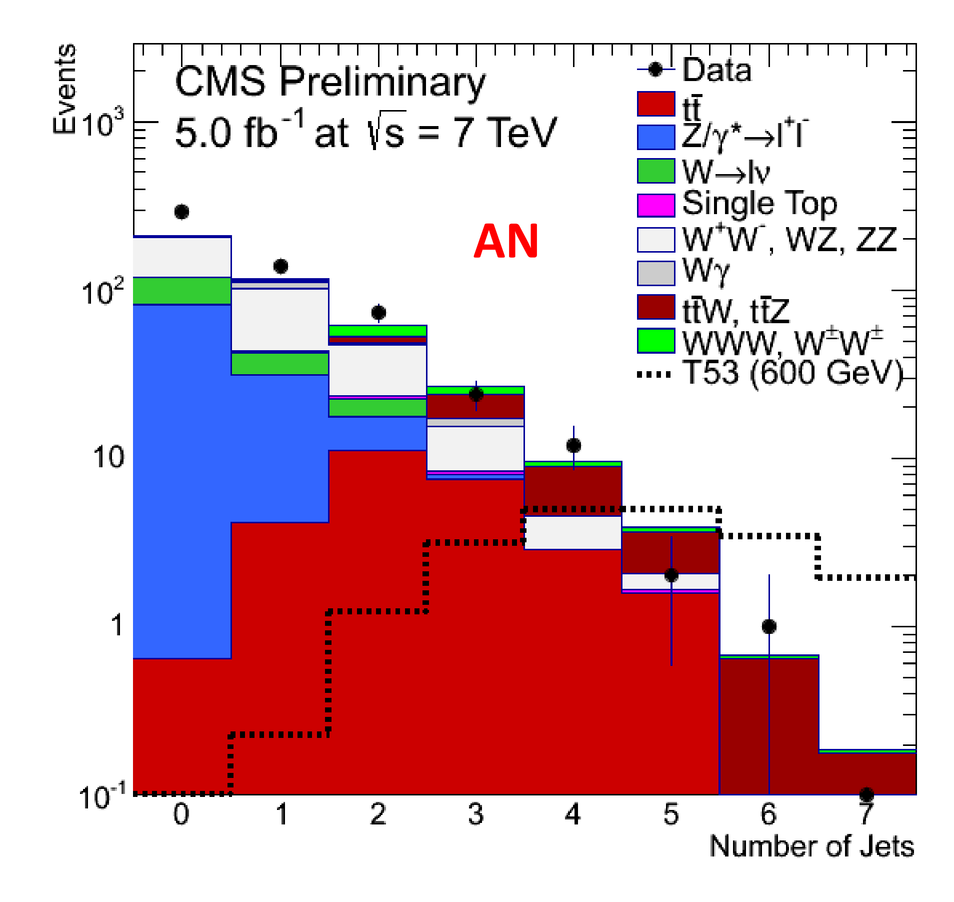
Object Reconstruction: Jets

- Jet selection (from TopLikeBSM twiki):
  - Anti-k_T particle flow jets
  - $p_T > 30$ GeV
  - $|\eta| < 2.4$
  - Charged hadron subtraction, L1FastJet corrections and L2L3 jet energy scale corrections
  - Loose particle flow jet ID
  - ΔR (selected leptons, jet) > 0.3

- The resolution of jets in MC is increased by 10% to account for data-MC difference

Jet p_T and nJets



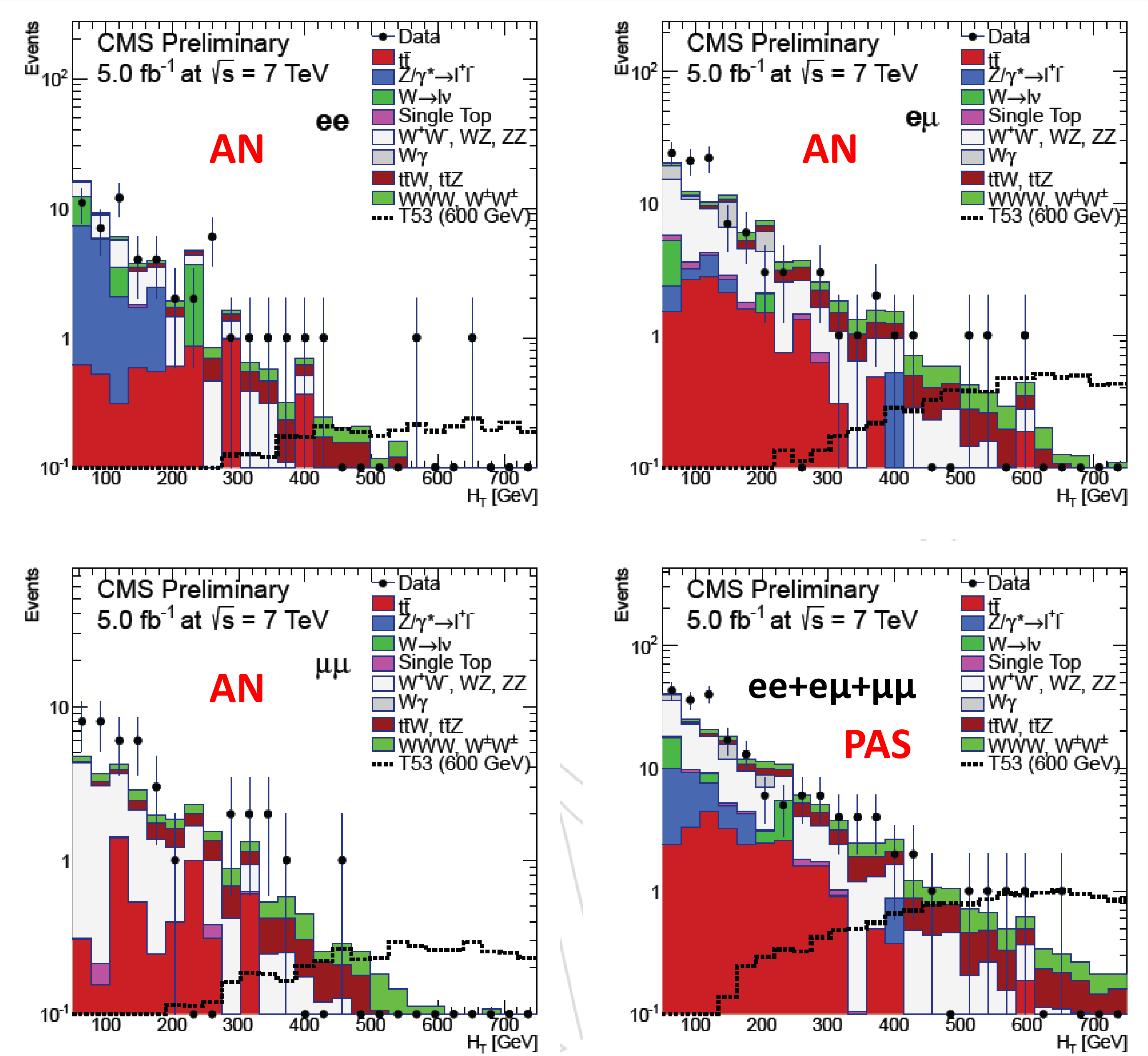


Leading jet p_T and number of jets after the same-sign lepton requirement and Z-veto (all three channels combined)

Final Event Selection

- Exactly two leptons
  - Reject events with 3 or more leptons
  - For the eμ channel, ΔR (electron, muon) > 0.1
- Leptons must have same charge
- Z and quarkonia invariant mass vetos:
  - M(ll) < 76 GeV or M(ll) > 106 GeV
  - M(ll) > 20 GeV
- At least 4 jets
- H_T = sum of selected jet p_T > 300 GeV
- Cuts optimized using expected limit
  - No significant difference with variation of lepton p_T, H_T

H_T after same-sign leptons and Z-veto



Signal Yields

$T_{5/3}$ Mass (GeV)	2SS leptons	$M(\ell\ell)$ Veto	N(jet)$\geq$ 4	$H_T \geq 300$
400	279	249	188	170 ± 1.99
450	138	126	97.0	91.1 ± 1.00
500	70.8	66.6	50.7	48.6 ± 0.51
550	37.7	35.9	28.1	27.1 ± 0.28
600	20.7	19.9	15.3	14.9 ± 0.15
650	11.6	11.3	8.71	8.55 ± 0.08
700	6.73	6.55	5.07	5.00 ± 0.05
PAS 750	3.98	3.90	2.99	2.96 ± 0.03

Table 8: Summary table of expected signal events in all three channels.

MC Background/Data Yields

PAS Sample	2SS leptons	$M(\ell\ell)$ Veto	N(jet)$\geq$ 4	$H_T \geq 300$
$t\bar{t}$	32.55	27.83	4.45	1.63 ± 0.68
Z+Jets	1171	115.7	0.00	0.00 ± 0.00
W+Jets	62.46	51.91	0.00	0.00 ± 0.00
$W^{\pm}W^{\mp}$, WZ, ZZ	200.6	177.54	2.19	1.40 ± 0.08
Single Top	3.22	2.69	0.07	0.00 ± 0.00
$W\gamma$	24.94	17.31	0.00	0.00 ± 0.00
$W^{\pm}W^{\pm}$	15.96	14.58	0.72	0.63 ± 0.06
WWW	2.36	2.11	0.10	0.06 ± 0.00
$t\bar{t}$ W	12.83	11.51	3.80	2.52 ± 0.05
$t\bar{t}$ Z	3.52	3.16	1.32	0.90 ± 0.03
Total MC	1529.21	424.37	12.64	7.14 ± 0.69
Data	1609	542	15	10

All three channels combined. The MC is only used to estimate final yields for processes outlined in red (others are derived using data-driven methods).

Background Determination

- From simulation
  - Rare Standard Model same-sign backgrounds
- Data-driven
  - Same sign backgrounds
    - Jets mis-identified as leptons
    - Non-prompt leptons from decays in flight or heavy flavors
  - Opposite-sign backgrounds
    - From charge mis-id

Same-Sign Non-Prompt Backgrounds

- Use Tight-Loose method
  - See AN-2010/261, AN-2010/257, AN-2011/258 for details
  - Define looser lepton criteria than are used in the analysis
  - Calculate the “fake rate” (FR): probability that a lepton passing the loose criteria also passes the tight ones
  - Estimate the number of events with non-prompt leptons passing the tight selection
- Fake rate weight: $\varepsilon_\ell = \frac{FR_\ell}{1-FR_\ell}$
- For ee, μμ:
$$N_{tf} = \varepsilon_\ell N_{t10} - 2\varepsilon_\ell^2 N_{t00}$$
$$N_{ff} = \varepsilon_\ell^2 N_{t00}$$
- For eμ:
$$N_{tf} = \varepsilon_{\mathrm{e}} N_{t01} + \varepsilon_\mu N_{t10} - 2\varepsilon_{\mathrm{e}}\varepsilon_\mu N_{t00}$$
$$N_{ff} = \varepsilon_{\mathrm{e}}\varepsilon_\mu N_{t00}$$

Loose Leptons

- Electrons
  - $p_T > 30$ GeV
  - $|\eta| < 2.4$ except EB-EE gap ($1.4442 < |\eta_{sc}| < 1.566$)
  - HyperTight1MC Cut in Categories electron ID
  - PF-based RelIso < 0.60 (vs. 0.15)
  - Transverse IP < 0.02 cm
  - Gsf-Ctf-ScPix charge consistency
  - No conversion rejection
- Muons
  - $p_T > 30$ GeV
  - $|\eta| < 2.4$
  - Global Muon and Tracker Muon
  - PF-based RelIso < 0.40 (vs. 0.2)
  - χ^2 / ndof < 50 (vs. 10)
  - Transverse IP < 0.02 cm
  - N Muon hits > 0
  - N Pixel hits > 0
  - No requirement on valid silicon hits (vs. > 10)
  - No requirement on chambers with matched segments (vs. > 1)

Computing the Fake Rates

- Use EleHad, MuHad datasets (Multijet for cross-check)

- Remove events with leptons from W/Z:
  - At least one jet
  - At least one loose lepton
  - Missing ET < 25 GeV
  - Transverse mass < 25 GeV
  - ΔR (lepton, jet) < 1
  - Z-veto: M(ll) < 76 GeV or M(ll) > 106 GeV

- Fake rate is then $FR = \frac{\text{tight leptons}}{\text{loose leptons}}$

- Muon FR: 0.453 +/- 0.002
- Electron FR: 0.333 +/- 0.012

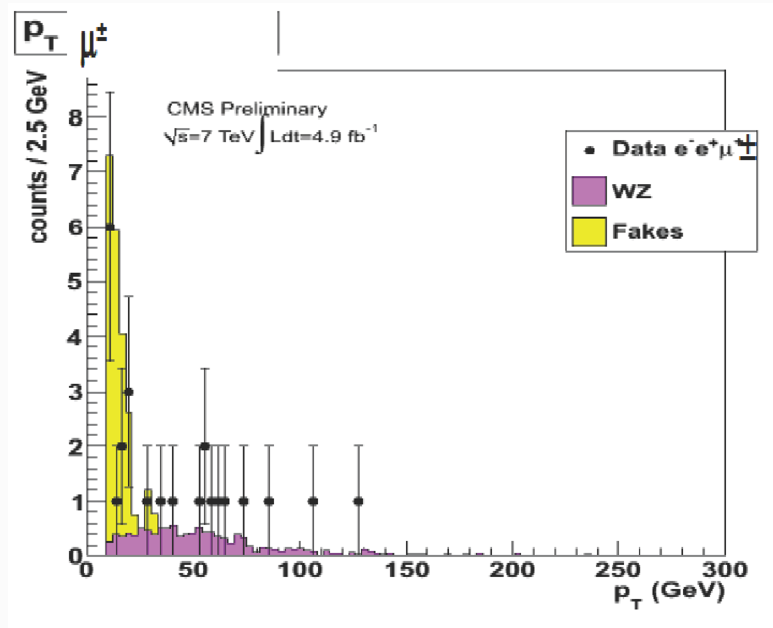
Non-Prompt Background Results

Selection	SS	Z_{veto}	N(jets)	H_T
PAS		$\mu\mu$ channel		
N_{t00}	38 ± 6.16	24 ± 4.89	0 ± 0.00	0 ± 0.00
N_{t10}	112 ± 10.58	74 ± 8.60	10 ± 3.1	6 ± 2.44
N_{ff}	26.14 ± 4.23	16.51 ± 3.36	0 ± 0.00	0 ± 0.00
N_{tf}	40.61 ± 9.69	28.36 ± 8.02	8.29 ± 2.57	4.97 ± 2.02
Nfakes	66.75 ± 10.57	14.92 ± 8.69	8.29 ± 2.57	$\boxed{4.97 \pm 2.02}$
		ee channel		
N_{t00}	1081 ± 32.87	599 ± 24.27	6 ± 2.44	4 ± 2.00
N_{t10}	2927 ± 54.10	642 ± 25.22	13 ± 3.51	6 ± 2.44
N_{ff}	271.12 ± 8.24	150.23 ± 6.08	1.50 ± 0.61	1.00 ± 0.50
N_{tf}	923.61 ± 27.69	21.05 ± 13.56	3.50 ± 2.35	0.99 ± 1.86
Nfakes	1194.83 ± 28.89	171.28 ± 14.86	5.00 ± 2.42	$\boxed{1.99 \pm 1.92}$
		$e\mu$ channel		
N_{t00}	298 ± 17.26	175 ± 13.22	1 ± 1.00	1 ± 1.00
N_{t10}	100 ± 10.00	61 ± 7.81	2 ± 1.41	0 ± 0.00
N_{t01}	828 ± 28.77	526 ± 22.93	15 ± 3.87	8 ± 2.82
N_{ff}	123.79 ± 17.26	72.69 ± 5.49	0.42 ± 0.42	0.42 ± 0.42
N_{tf}	250.03 ± 44.33	168.64 ± 24.66	8.33 ± 3.87	3.16 ± 2.58
Nfakes	373.82 ± 47.57	241.33 ± 25.26	8.75 ± 3.89	$\boxed{3.58 \pm 2.61}$

Closure Tests

- Check to see that the method can correctly describe events with multiple leptons that have an extra non-prompt lepton
  - Use events with leptonically decaying Z bosons that have an extra lepton (81 GeV < M(ll) < 101 GeV)
  - Reduce WZ contribution by requiring missing E_T < 20 GeV
  - Subtract MC estimate of remaining WZ and Vγ

- $e^+e^-\mu$:
  - With fake rate: 14.6 +/- 2.0
  - Expected: 14.0 +/- 0.5
- $\mu^+\mu^-e$:
  - With fake rate: 20.0 +/- 3.3
  - Expected: 18.3 +/ 6.7



Closure Tests

- Consider Monte-Carlo W+Jets events and ask for two leptons
  - Expect over-estimate since fake rate in MC (as measured on QCD events) is smaller than that for data
- Check the Multijet dataset and look at events with same-signed leptons only

Sample	Event type	Observed	Expected
W+jets (MC)	2 SS Ee	31	52.83 ± 7.00
W+jets (MC)	2 SS $\mu\mu$	0	1.12 ± 1.53
W+jets (MC)	2 SS eμ	34	60.54 ± 9.13
multi-jet (data)	2 SS eμ	29	47.1 ± 8.0
multi-jet (data)	2 SS ee + $MET < 25$ + Z mass cut	10	8.8 ± 3.91

Table 18: Summary of various closure tests on data and Monte Carlo samples.

- Also considered parameterizations based on p_T and η
  - Difference is small compared to statistical uncertainty
- Assign 50% systematic uncertainty to cover variation seen in closure tests and pT and η parameterizations

Lepton Charge Mis-ID

- Expect net effect of charge misID to be small
  - Negligible for muons (probability of 10^{-4} to 10^{-5})
    - See CMS-SUS-020-11, arXiv:1205.3933
- Consider events with 2 electrons Z mass window
  - Apply cleaning, trigger and lepton quality requirements
    - No jet or H_T requirements

PAS	Total Bckd	$t\bar{t}$	Z+jets	W+jets	WZ	ZZ	$t\bar{t}$ W	$t\bar{t}$ Z
In Z-Window	9.3e+05	644	9.3e+05	35	205	46.2	1.89	8.62
Same-sign	1090	1.84	1050	9.97	8.9	1.02	0.62	0.20

Dominated by Z+jets

Electron Charge Mis-ID

- Compute probability of charge mis-ID in data: 5.9E-4
- Apply full event selection (including jets and H_T), but with oppositely charged leptons

	Opposite-sign	Expected same-sign
ee	171	0.20 ± 0.02
$e\mu$	308	0.18 ± 0.01

PAS

- Method checked using calorimeter electrons only
  - No noticeable difference (less than 1%)
- 20% systematic uncertainty

Systematics

- Luminosity (2.2%)
- Jets and pileup uncertainties obtained by varying the jets and pileup respectively according to the official recipes
  - For signal, 1% for JES and 2% for pileup
- Trigger uncertainties from AN-11-333
- Lepton efficiency from EXO-12-001
- PDF uncertainty of 1.5%

Effect	Uncertainty (%)
Electron trigger	0.6
Muon trigger	0.5
Lepton efficiency	3.0
Luminosity	2.2

PAS

Sample	JES	Pileup	Normalization
WZ	5.0%	1.8%	17%
ZZ	1.1%	0.6%	7.5%
$W^{\pm}W^{\pm}$	4.5%	2.4%	50%
WWW	3.7%	0.5%	50%
$t\bar{t}W$	3.4%	0.94%	50%
$t\bar{t}Z$	3.7%	0.25%	50%

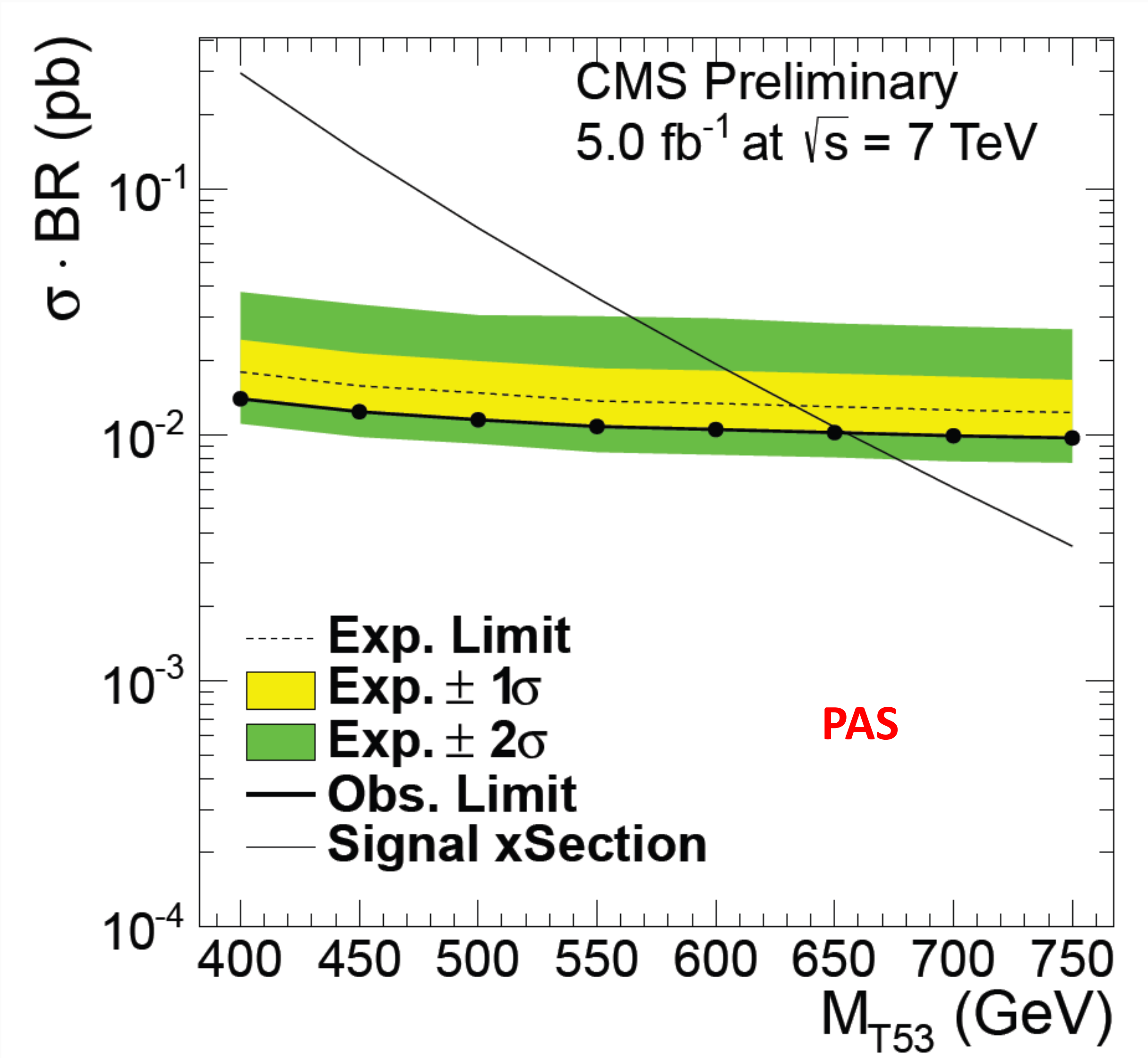
PAS

Results

PAS	PSS MC	Fake Leptons	Charge Mis-ID	Total Expected	Observed
ee	1.1 ± 0.6	2.0 ± 2.2	0.2 ± 0.04	3.3 ± 1.9	2
$e\mu$	3.0 ± 1.6	3.6 ± 3.2	0.2 ± 0.04	6.8 ± 2.6	6
$\mu\mu$	1.3 ± 0.7	5.0 ± 3.2	-	6.3 ± 2.0	2
All	5.5 ± 3.2	10.5 ± 6.5	0.4 ± 0.1	16.4 ± 7.2	10

- Summary of data and backgrounds from various sources
  - PSS MC are the irreducible, prompt same sign backgrounds
- Dominated by instrumental background from fake leptons
- Includes all systematic uncertainties

Limit



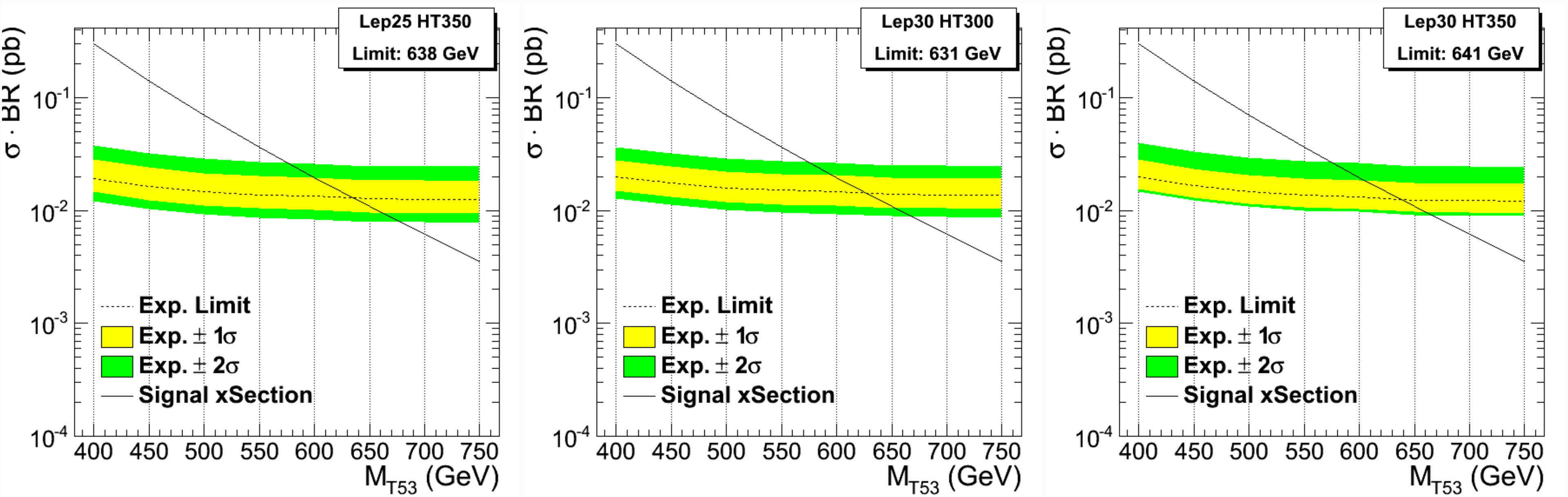
- Limit Evaluated with ROOSTATs
- CLS method
- Expected limit: 635 GeV
- Observed limit: 655 GeV

Summary

- We have searched for an exotic top partner with charge 5/3e
  - Same-sign dilepton ee, eμ, μμ channels
- No significant excess observed with 5.0 fb^{-1} at 7 TeV
- Exclude at 95% C.L. masses of such particles below 655 GeV
  - 635 GeV expected

Backup

Cut Optimization



Fake Rate Method

- Use Tight-Loose method
  - See AN-2010/261, AN-2010/257, AN-2011/258

- Define loose selection for each kind of lepton
  - Let N_l = number of dilepton events passing loose selection
  - Let N_{tx} = number of such events where x leptons also pass the tight selection
  - p = probability of prompt loose lepton to be tight
  - f or FR = probability of non-prompt loose lepton to be tight
  - N_{pp}, N_{pf}, N_{ff}: Two prompt, one prompt, both fake respectively

AN-2010/261

$$
\begin{aligned}
N_l &= N_{pp} + N_{fp} + N_{ff} = N_{t2} + N_{t1} + N_{t0} \\
N_{t0} &= (1-p)^2 N_{pp} + (1-p)(1-f) N_{fp} + (1-f)^2 N_{ff} \\
N_{t1} &= 2p(1-p) N_{pp} + [f(1-p) + p(1-f)]\, N_{fp} + 2f(1-f) N_{ff} \\
N_{t2} &= p^2 N_{pp} + pf N_{fp} + f^2 N_{ff}
\end{aligned}
$$

Fake Rate Method

- Assumptions:
  - Probability of a lepton to be non-prompt is independent of the presence of another lepton
  - p = 1 (all prompt loose leptons are also tight)
    - It's about 95% in Monte-Carlo

- Solve equations for N_{pf} (also called N_{tf}) and N_{ff}
  - Define fake rate weight $\varepsilon_\ell = \frac{FR_\ell}{1-FR_\ell}$
  - Then

$$N_{tf} = \varepsilon_\ell N_{t10} - 2\varepsilon_\ell^2 N_{t00}$$
$$N_{ff} = \varepsilon_\ell^2 N_{t00}$$

  - For the eμ channel:

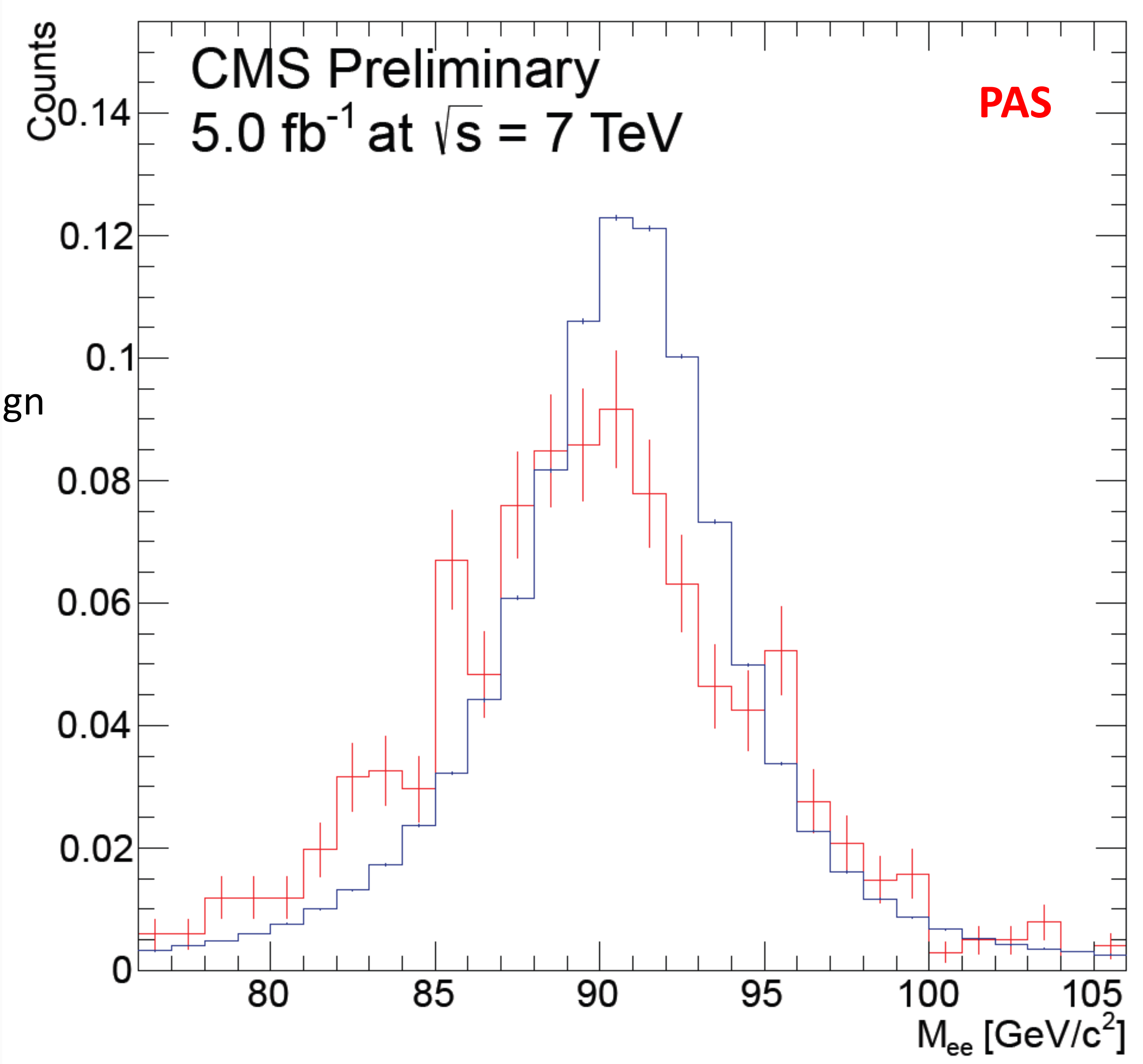
$$N_{tf} = \varepsilon_{\mathrm{e}} N_{t01} + \varepsilon_\mu N_{t10} - 2\varepsilon_{\mathrm{e}}\varepsilon_\mu N_{t00}$$
$$N_{ff} = \varepsilon_{\mathrm{e}}\varepsilon_\mu N_{t00}$$

Fake Rate: TTbar Closure Tests

	ee	eμ	μμ
Predicted	11.62	6.57	14.91
Observed	16.22	4.60	10.52

Charge MisID Invariant Mass Comparison

Red = Same Sign
Blue = Opposite Sign



Bayesian Limit