Same-sign and Multi-lepton signatures

Xin Chen, Bruce Mellado, Yibin Pan

University of Wisconsin–Madison

Characterization of New Physics at the LHC II
CERN, Nov. 5-6, 2010

The topology-based search

- There are quite a few topology sets, based on photons/leptons/jets/MET multiplicity/hardness, defined by theorists of the LHC New Physics Working Group. We investigated a few of them involving same-sign (SS) dilepton or multilepton topologies (there are more relevant models not covered in the talk)

- We want to set up the mapping between the theoretical (hadron level MC truth) and experimental (simulated using MadGraph PGS and/or simple smearing of particle energy with nominal calorimeter resolution, see backup slides) worlds, and understand the simulation (detector acceptance, trigger, offline ID/cuts) efficiency for each topology set

- We make fits to pseudodata sets to test the sensitivity for each topology as a function of σxBR and integrated luminosity, assuming that the signal efficiency is roughly stable for different mass parameter values in the model

- We try to minimize the reliance on MC predictions by making use of side band (*SB*) events and/or control samples (*CS*) from data to make model-parameter-independent statistical interpretations

The same-sign and multi-lepton topologies

The SS dilepton and multi-lepton topologies are the hot beds for New Physics due to the smallness of the SM background with this signature. We studied four topology sets in this talk:

1) Majorana neutrinos in the LRSM: $q\bar{q}\to W_R^{\pm}\to l^{\pm}N_R(N_R\to l^{\pm}jj)$. Events generated with Pythia and with $m(W_R^{\pm})=1200\,\mathrm{GeV}$, $m(N_R)=600\,\mathrm{GeV}$. Nominal $\sigma\times BR$ is 0.6 pb^{-1}. Topology characterized by 2 SS leptons, 2 jets and no MET

2) Squark-pair-like SSL process: $qq\to\tilde{q}\tilde{q}(\tilde{q}\to C_1^{\pm}j, C_1^{\pm}\to W^{\pm}N_1)$. Events generated with MadGraph+Pythia (many thanks to *Veronica Sanz* for the recipe of MadGraph+BRIDGE) and with $m(\tilde{q})=500\,\mathrm{GeV}$, $m(C_1)=200\,\mathrm{GeV}$, $m(N_1)=100\,\mathrm{GeV}$. Topology characterized by 2 SS leptons, 2 jets and large MET

3) Gluino-pair-like SSL process: $q\bar{q}, gg\to\tilde{g}\tilde{g}(\tilde{g}\to C_1^{\pm}jj, C_1^{\pm}\to W^{\pm}N_1)$. Events generated with MadGraph+Pythia (scripts/cards by *P. Schuster, N. Toro, J. Wacker* and *E. Izaguirre*) and with $m(\tilde{g})=400\,\mathrm{GeV}$, $m(C_1)=250\,\mathrm{GeV}$, $m(N_1)=100\,\mathrm{GeV}$ Topology characterized by 2 SS leptons, 4 jets and large MET

4) Electroweakino-pair process: $q\bar{q}\to W^{\pm *}\to C_1^{\pm}N_2(C_1^{\pm}\to W^{\pm}N_1, N_2\to l^{+}l^{-}N_1)$. Events generated with MadGraph+Pythia (scripts/cards by *T. Tait, J. Wacker, P. Schuster, M. Lisanti* and *J. Kaplan*) and with $m(C_1^{\pm})=250\,\mathrm{GeV}$, $m(N_2)=150\,\mathrm{GeV}$, $m(N_1)=100\,\mathrm{GeV}$. Topology characterized by 3 leptons, 0 jet and large MET

The main SM backgrounds for each topology

- The overall SM background for the SS 2-lepton and trilepton signals is small, but differences exist for the 4 topologies. The background rates, which are estimated from hadron-level particles, are rough estimates only. We need the side band and/or control sample to estimate the background rates in-situ

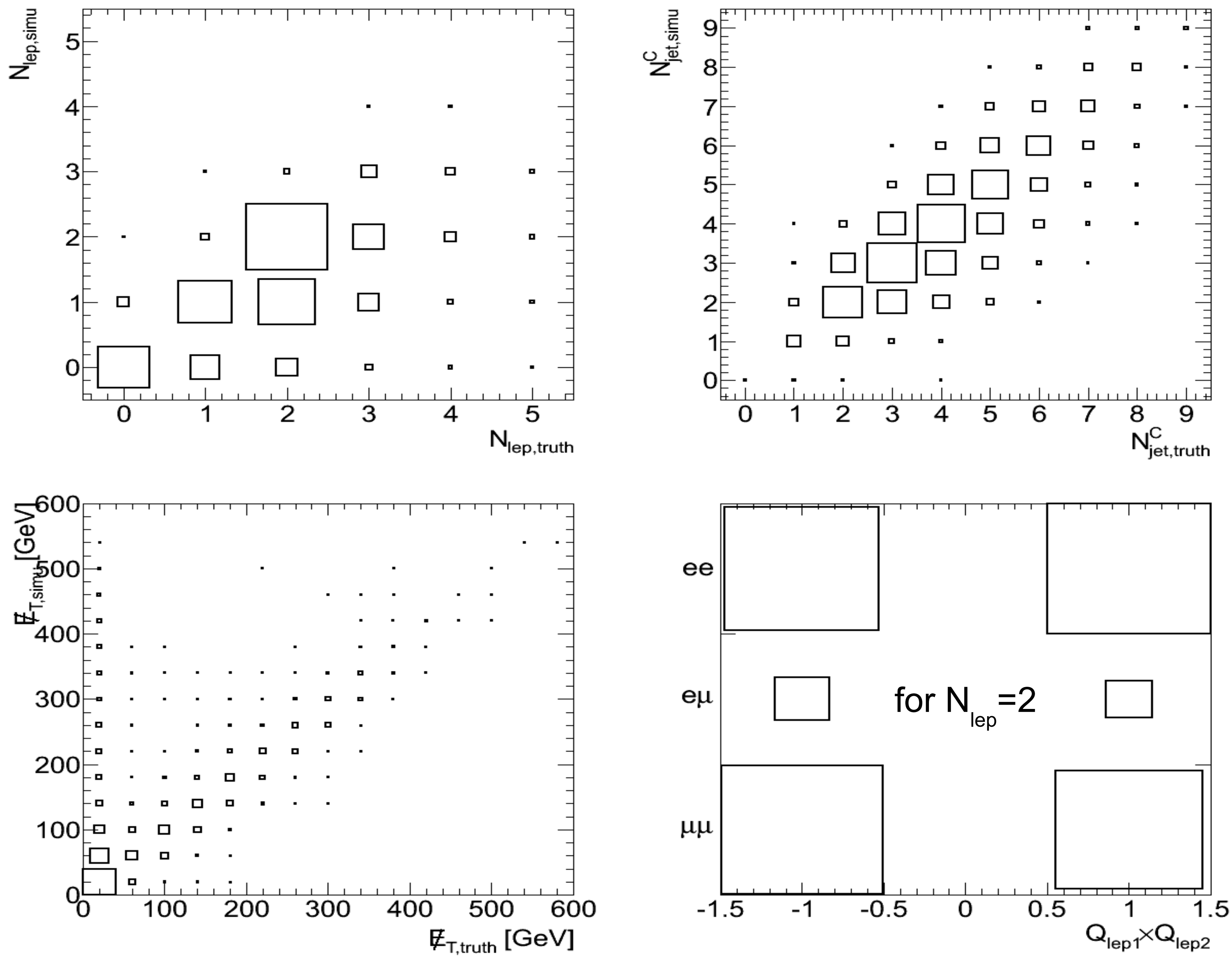
Topology (1): the main backgrounds are dijet, ttbar, W+jets and Z+jets. One of the 2 SS leptons (usually the softer one) in dijet, ttbar and W+jets mostly comes from jet fragmentation or jets faking electrons, whereas 2 SS leptons can be found in $Z \to ee$ due to electron charge misidentification rate (the fraction is on the order of $\sim 10^{-3}$–10^{-2}), which can be much reduced by a $Z \to ee$ mass veto

Topology (2) and (3): the dominant backgrounds are the same as for (1), but the QCD dijet and Z+jets can be killed by requiring a MET cut

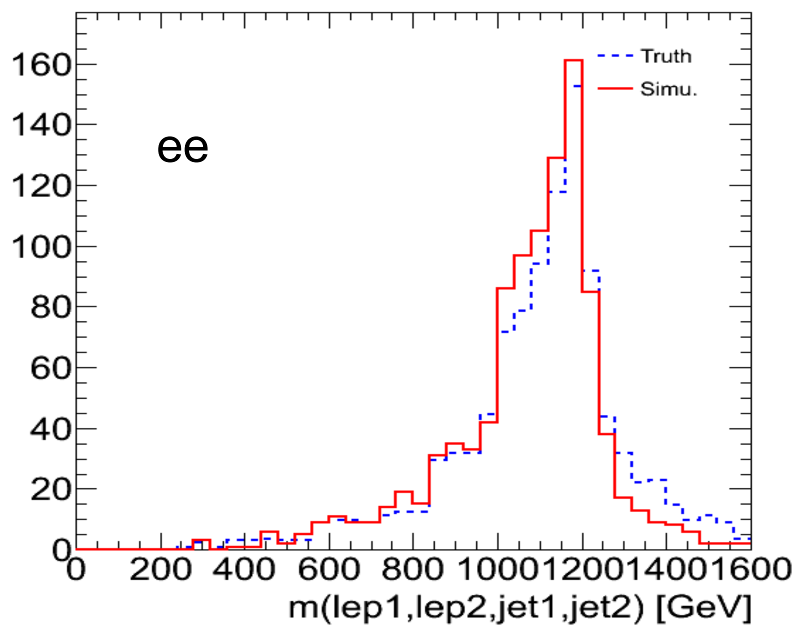
Topology (4): the dominant backgrounds are Z/γ^* (one of the trilepton is from jet fragmentation) and diboson (all 3 leptons from boson decays) after the central jet veto. An estimate of diboson WZ background is crucial

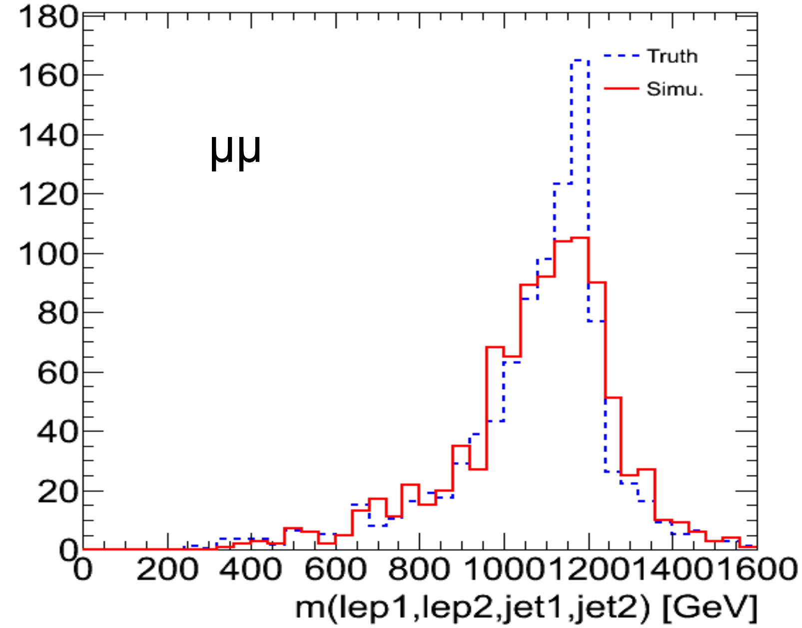
Topo (1): LRSM Majorana neutrino with 2 SS leptons, 2 jets and no MET
Topo (2): Squark-pair production with 2 SS leptons, 2 jets and large MET
Topo (3): Gluino-pair production with 2 SS leptons, 4 jets and large MET
Topo (4): Electroweakino-pair with 3 leptons, 0 jet and large MET

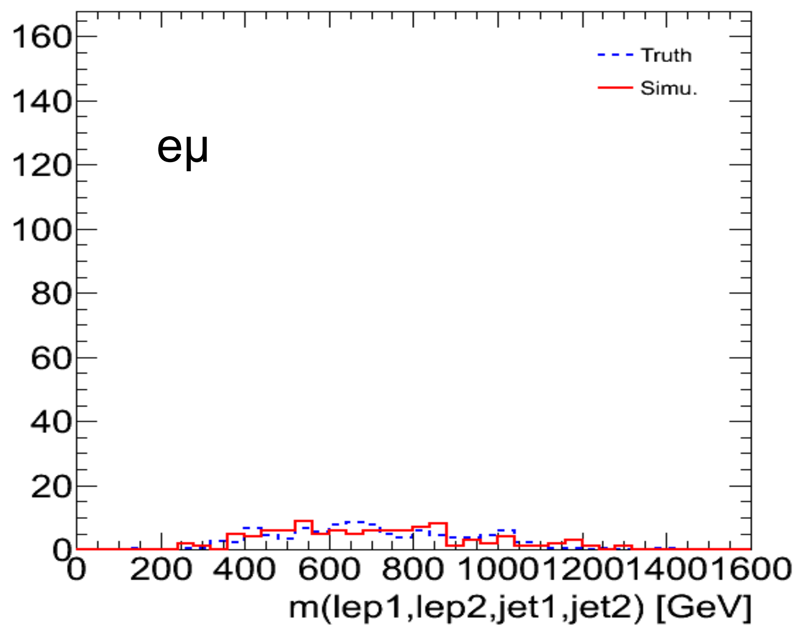
The LRSM topology



The LRSM topology – flavor composition

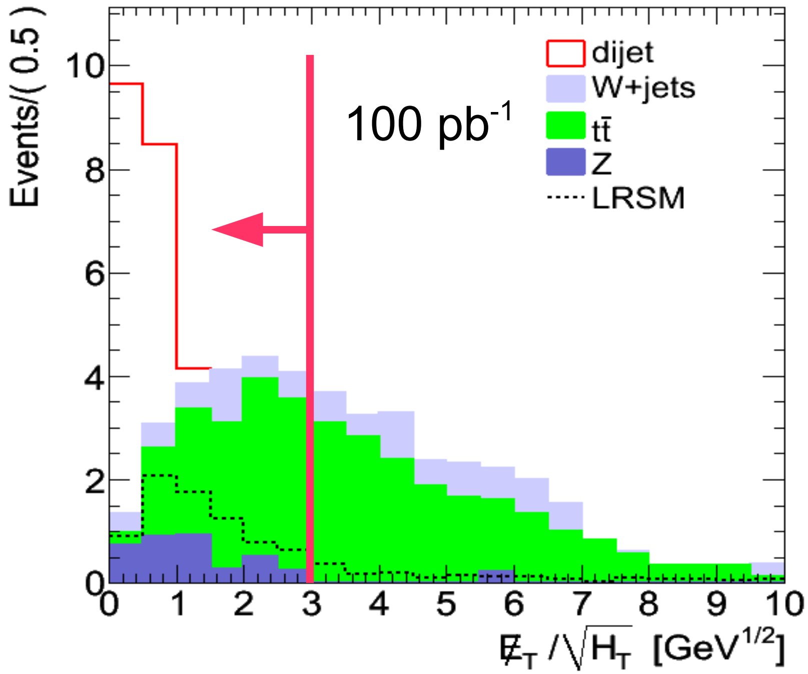


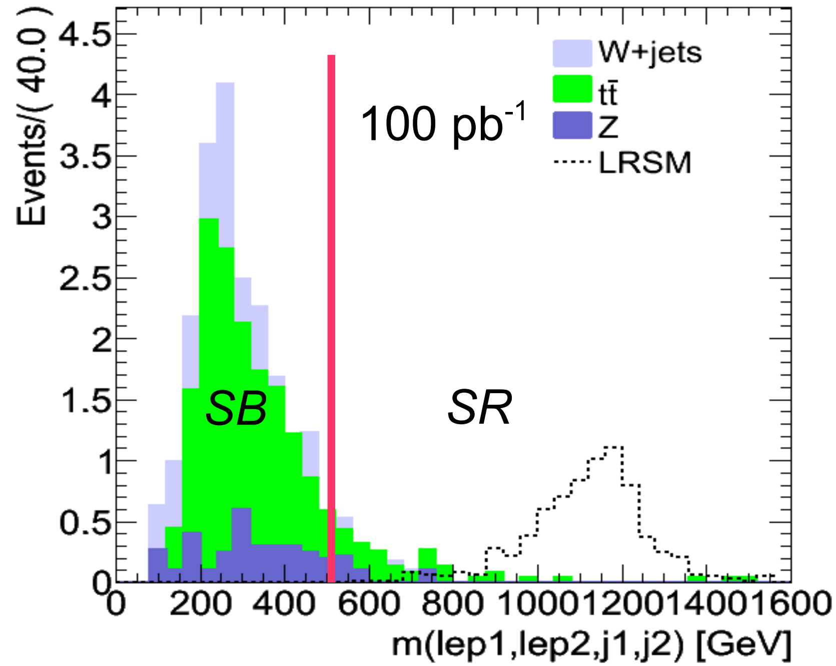




- No flavor mixing in the generation, roughly equal numbers of ee and μμ states are reconstructed
- No flavor separation is made in this talk. Later on we will subdivide in both flavor and charge in search for NP in all topologies

The LRSM signal and background





- The backgrounds are prepared by smearing the hadron-level truth particles of each process

- Require 2 SS leptons with p_T>30,10 GeV, veto of Z→ee events in the Z mass window, 2 central jets[1], $MET/\sqrt{H_T}$[2] <3 (*left plot*). The invariant mass of the 2 leading leptons and the 2 leading jets, after all the cuts, is shown in the *right plot* (dijet is not shown due to lack of MC statistics)

1) jets with p_T>20 GeV and $|\eta|$<2.5

2) H_T is the p_T sum of all central jets

Sensitivity to the LRSM topology

- Define m(lep1,lep2,j1,j2) > 500GeV as the signal-like region (*SR*), and the side band region (*SB*) is m(lep1,lep2,j1,j2) < 500GeV. The overall likelihood is

$$L=\frac{(s\mu+b\lambda)^{n_1}e^{-(s\mu+b\lambda)}}{n_1!}\frac{b^{n_2}e^{-b}}{n_2!}C_\lambda e^{\frac{-(\lambda-\lambda_{MC})^2}{2\Delta_\lambda^2}}C_\mu e^{\frac{-(\mu-1)^2}{2\Delta_\mu^2}} \quad (1)$$

n_1, n_2 are the total numbers of observed events in *SR* and *SB*

s, b are the expected numbers of signal and background in *SR* and *SB*

λ is the bg. extrapolation factor from *SB* to *SR*, Δ_λ the associated error

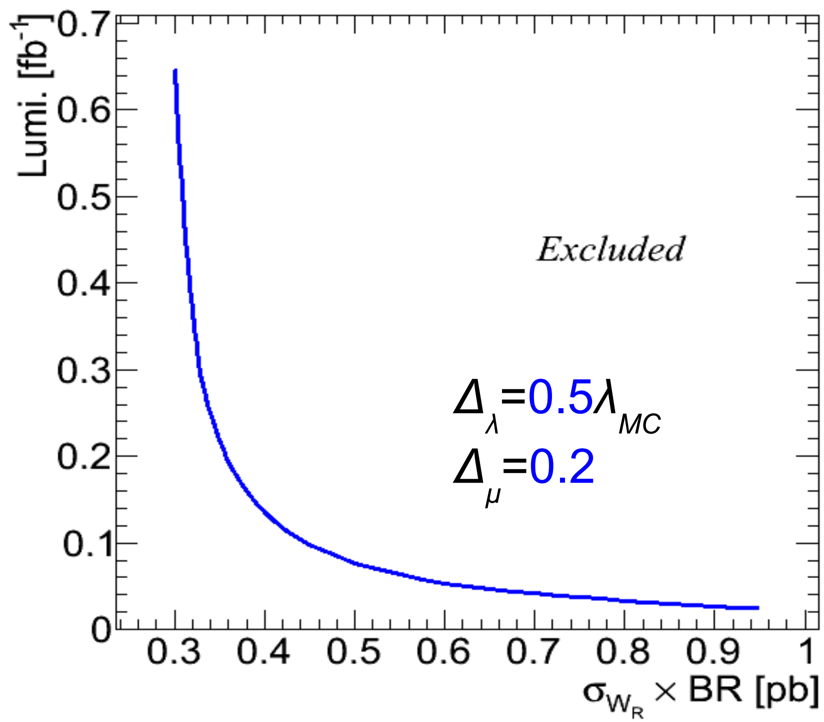
Δ_μ is the fractional error on the expected signal

The expected exclusion is plotted on the right using the log likelihood ratio:

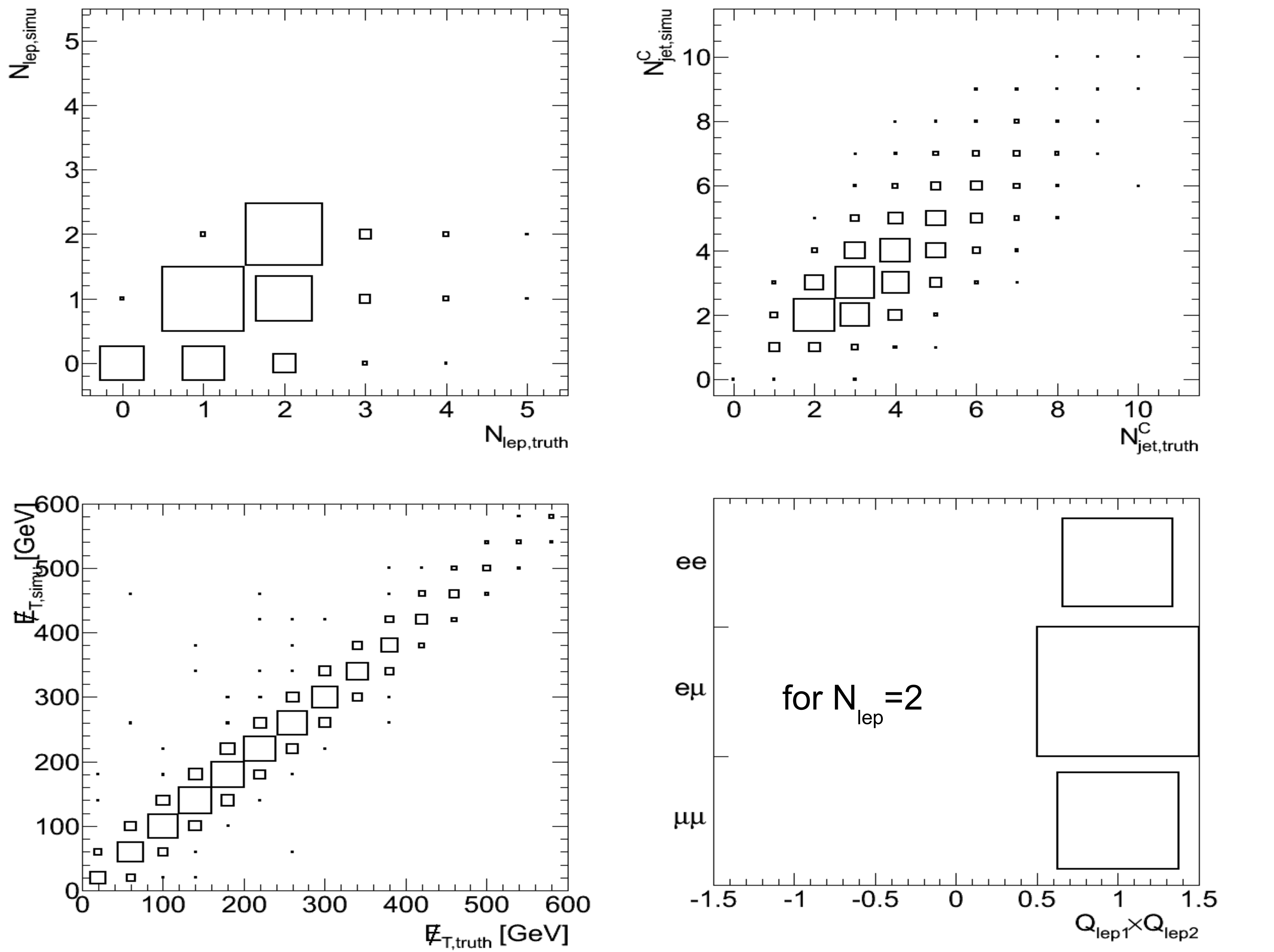
$$Z_{\text{excl}}=\sqrt{2\ln\frac{L_{\max}(s<s_{MC})}{L_{\max}(s=s_{MC})}}>1.645 \quad (2)$$

exclusion significance

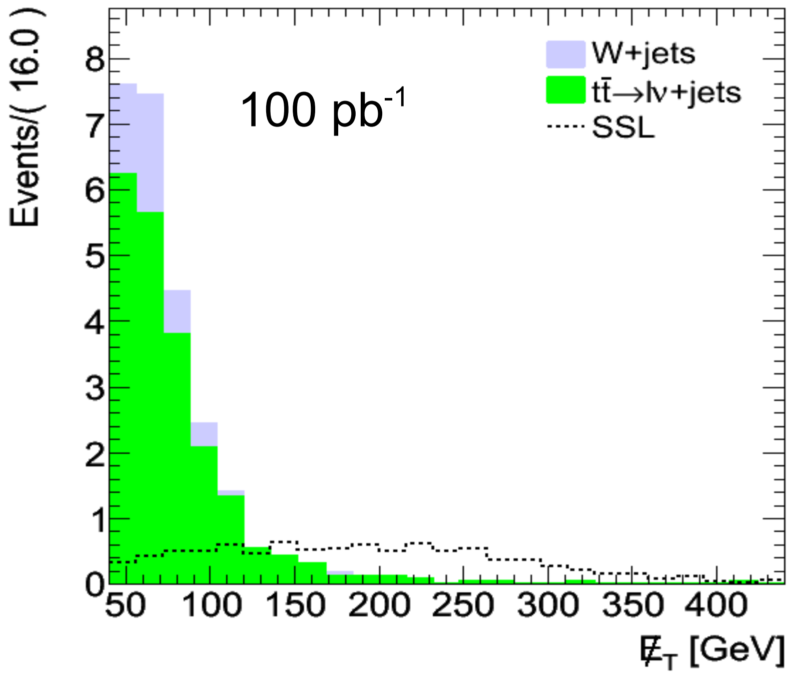
95% CL

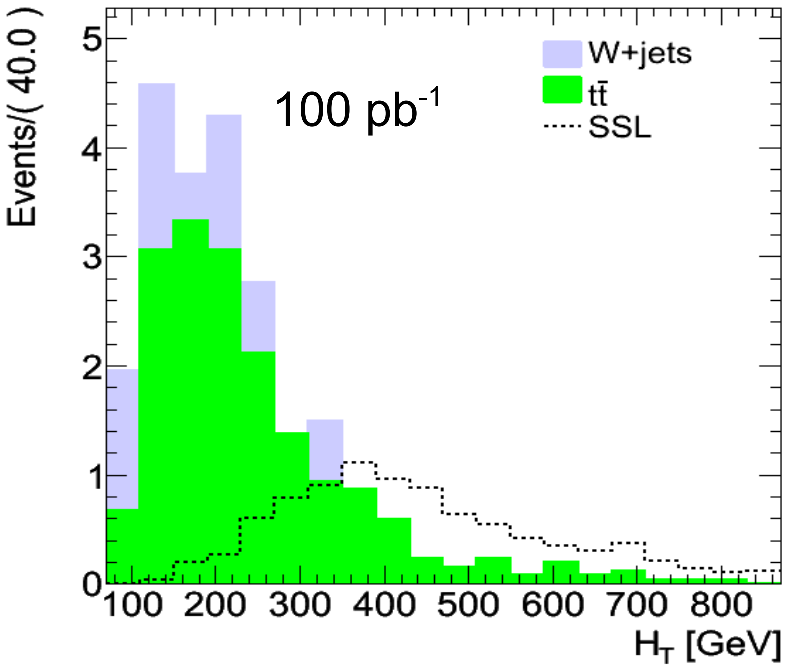


The Squark-pair topology



The Squark-pair signal and background

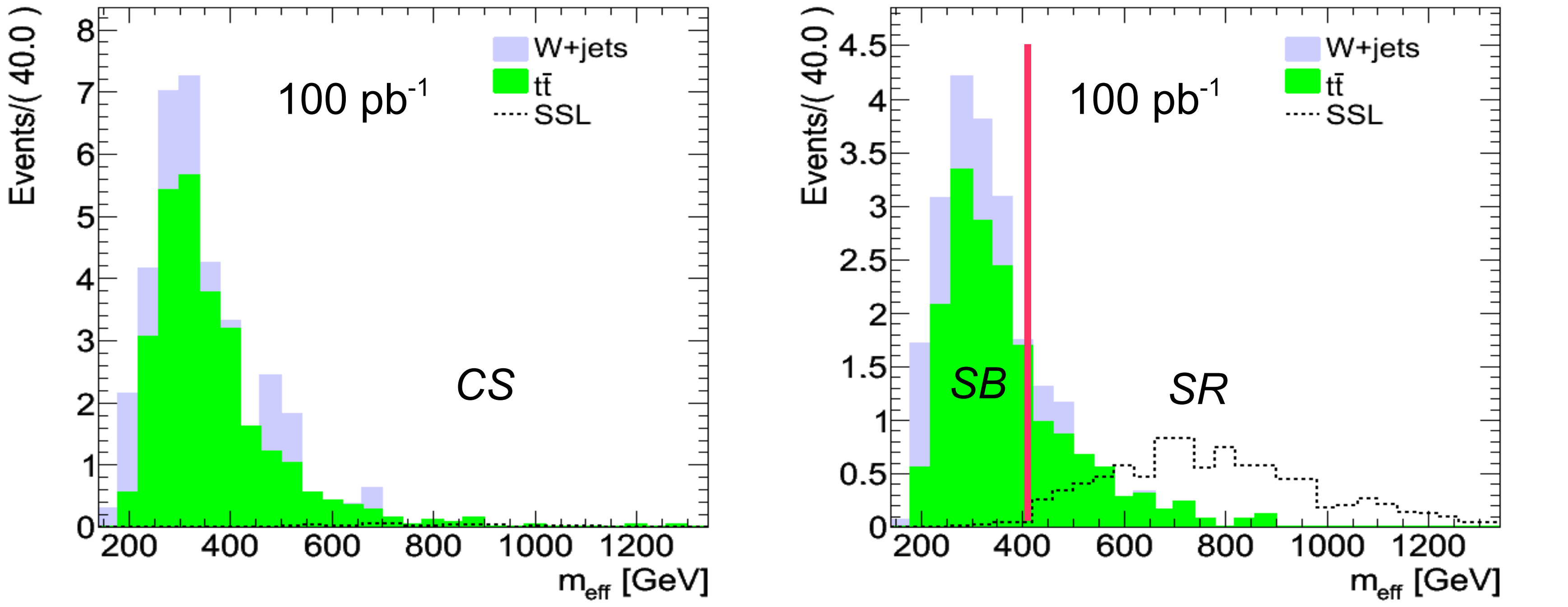




- Require 2 SS leptons with p_T>20,10 GeV, Z→ee veto, 2 central jets with p_T>40,30 GeV, MET>40GeV. The MET and H_T are plotted above

- Since the bg. is dominated by ttbar and W+jets (with one lepton from jet fragmentation), we can make a control sample (*CS*) by requiring one of the two leptons is a muon inside a jet (non-isolated) in the event, while all the other cuts stay the same (see next slide)

The Squark-pair signal and background



- $m_{eff}=p_T(lep1)+p_T(lep2)+MET+H_T$ distributions in the *CS* (*left plot*) and in the normal sample (*right plot*). The signal contamination in the *CS* is very small, and the background shape in the normal sample can be deduced from the *CS*

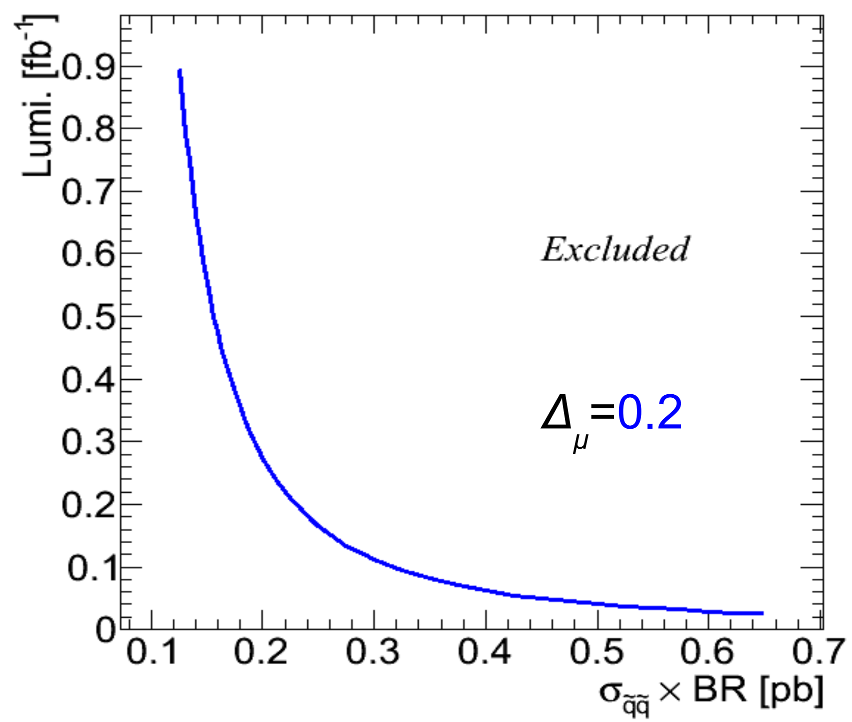
Sensitivity to the Squark-pair topology

- Define m_{eff} > 400GeV as the signal-like region (*SR*), and the side band region (*SB*) is m_{eff} < 400GeV. The overall likelihood is

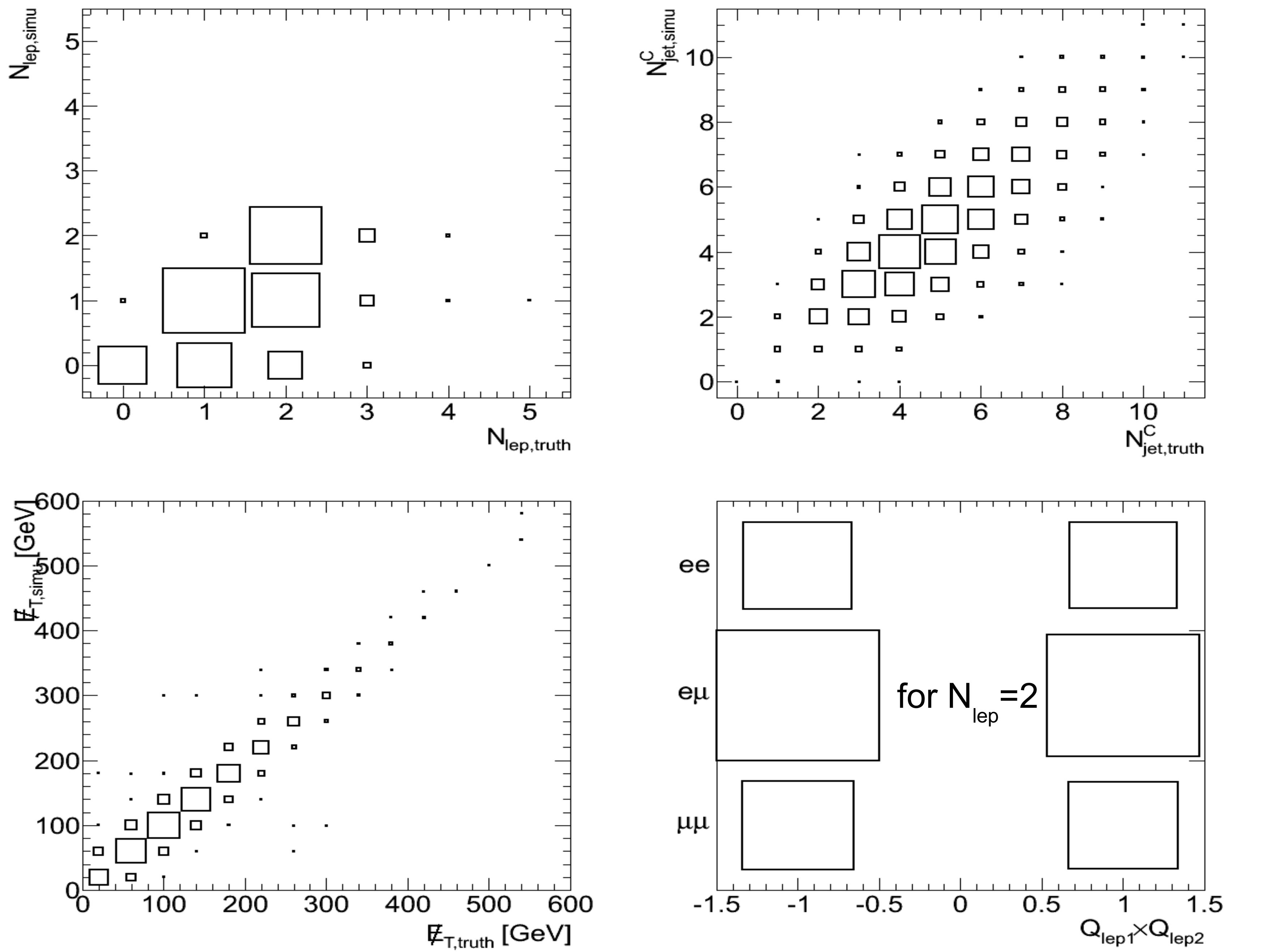
$$L=\frac{(s\mu+b\lambda)^{n_1}e^{-(s\mu+b\lambda)}}{n_1!}\frac{b^{n_2}e^{-b}}{n_2!}\frac{(n_3+n_4)!}{n_3!n_4!}\frac{\lambda^{n_3}}{(1+\lambda)^{n_3+n_4}}C_\mu e^{\frac{-(\mu-1)^2}{2\Delta_\mu^2}} \quad (3)$$

Definitions of variables are the same as in Eq. (1), except that we have a subsidiary measurement of *λ* in the *CS*, which gives a binormial term in the likelihood (n_3, n_4 are the *CS* counterparts of n_1 and n_2, repectively)

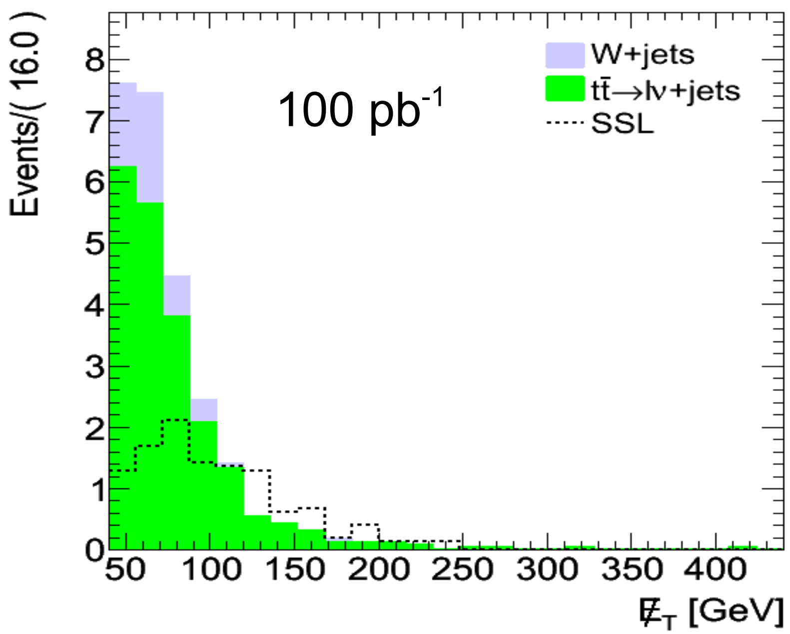
The expected exclusion is plotted on the right using the log likelihood ratio (Eq. 2)

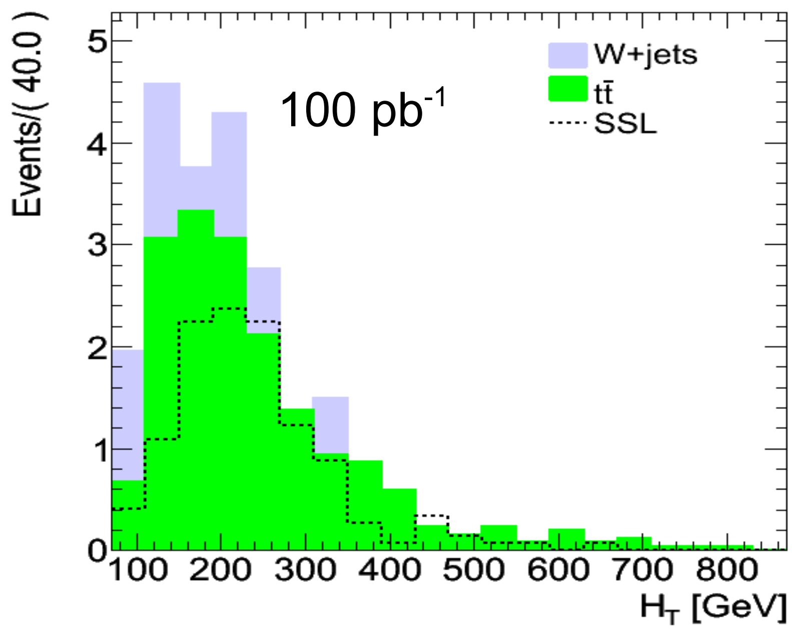


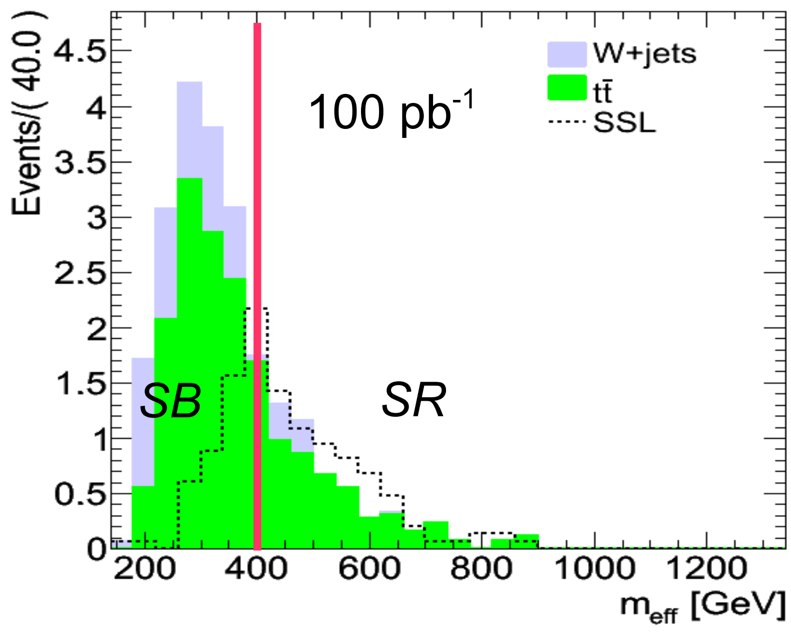
The Gluino-pair topology



The Gluino-pair signal and background



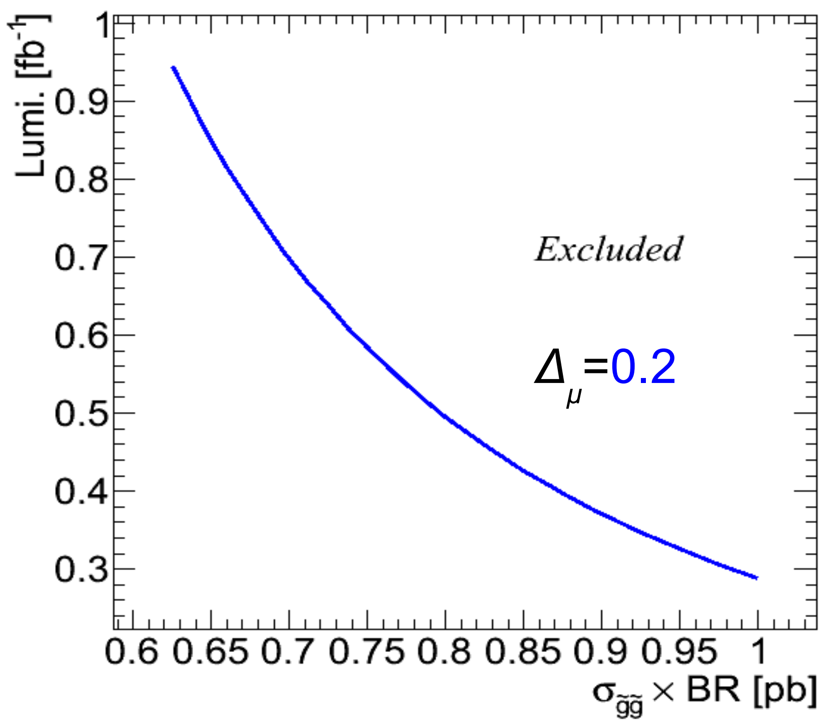




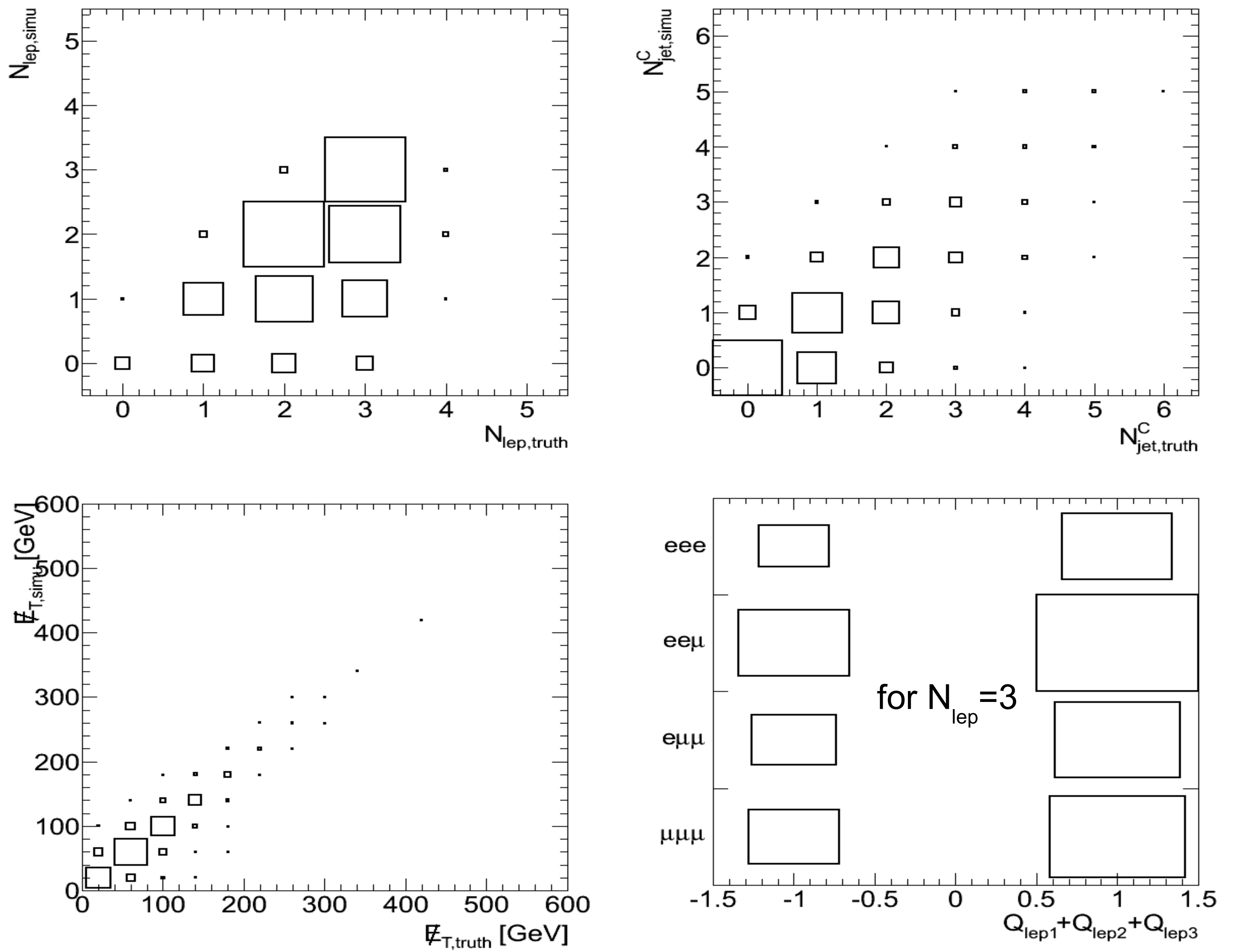
Reminder –
SR (signal region): m_{eff}>400GeV
SB (side band): m_{eff}<400 GeV

Sensitivity to the Gluino-pair topology

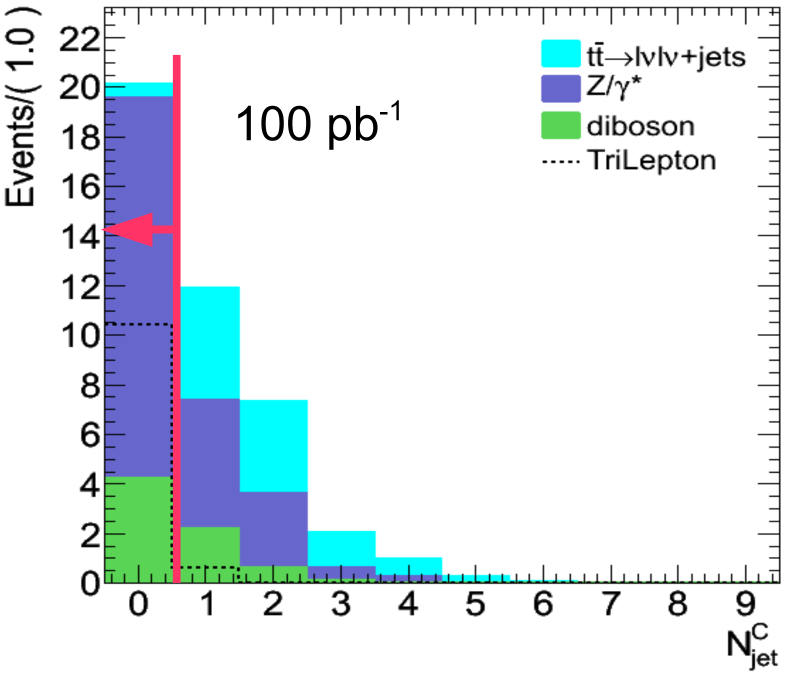
- Use the same definition of *SR*, *SB*, *CS* and likelihood (Eq. 3) as in the squark-pair topology, the expected exclusion is shown below
- The expected exclusion is plotted below using the log likelihood ratio (Eq. 2)

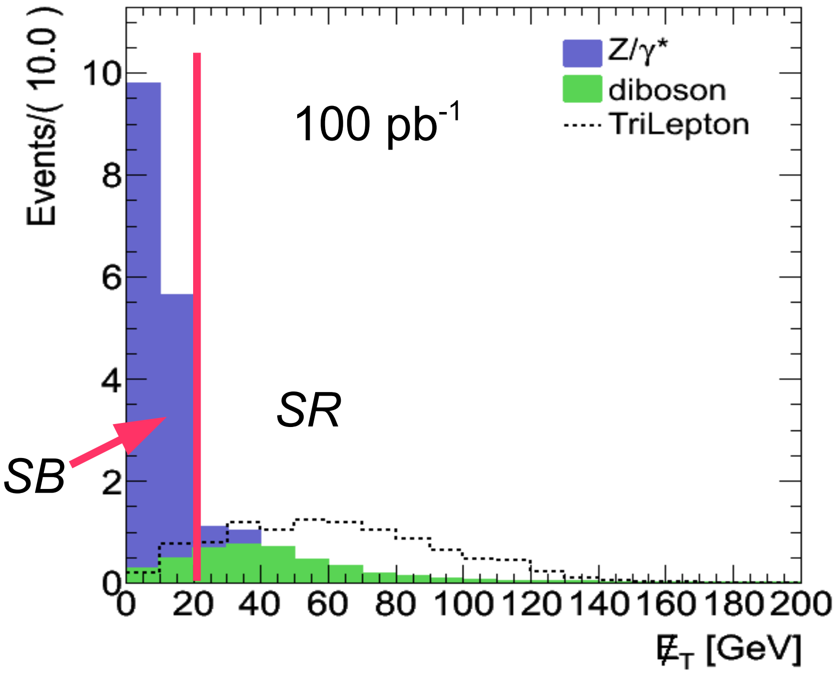


The Electroweakino-pair topology



The Electroweakino-pair signal and background





- Require 3 leptons with charges ++- or --+. The background is dominated by ttbar, Z/γ* and diboson WZ
- The jet activity in signal is low (*left plot*). The ttbar can be mostly reduced by a central jet veto ($N^{C}_{jet}=0$)
- The MET distribution after the central jet veto is shown in the *right plot*

Sensitivity to the Electroweakino-pair topology

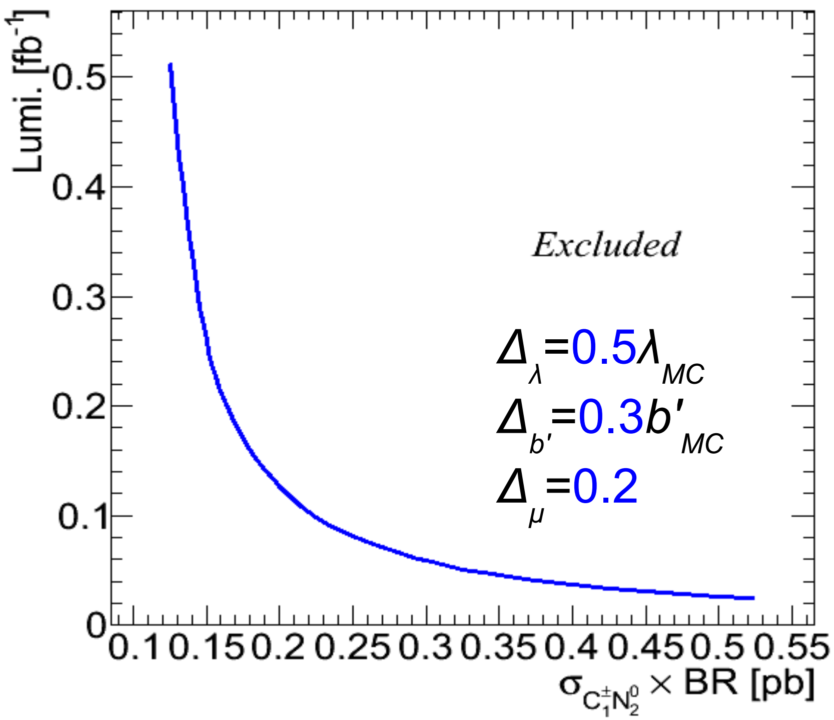
- Define MET > 20GeV as the *SR*, MET < 20GeV as the *SB*. The overall likelihood is

$$L=\frac{(s\mu+b\lambda+b')^{n_1}e^{-(s\mu+b\lambda+b')}}{n_1!}\frac{b^{n_2}e^{-b}}{n_2!}C_\lambda e^{\frac{-(\lambda-\lambda_{MC})^2}{2\Delta_\lambda^2}}C_{b'}e^{\frac{-(b'-b'_{MC})^2}{2\Delta_{b'}^2}}C_\mu e^{\frac{-(\mu-1)^2}{2\Delta_\mu^2}} \quad (4)$$

b represents the Z/γ* background in *SB*, and *b′* the diboson WZ in *SR*

The WZ rate is crucial for the triplepton analysis. It can be estimated by the WZ/Z* ratio from theory prediction, while the Z* rate in the relevant mass range can be well measured (next slide)

The expected exclusion is plotted on the right using the log likelihood ratio (Eq. 2)



Ratio ZW/Z$^{(*)}$

J. Campbell, N. Kauer, B. Mellado et al. at Phys. Rev. D80 : 054023, 2009

- **Ratio evaluated to NLO**
  - **For ZW require P_T or leading lepton to be >20 and the two sub-leading, >10 GeV**
  - **\sqrt{s} dependence of ratios evaluated**

$$R = \frac{\sigma^{NLO}_{q\bar{q}\to ZW}}{\sigma^{NLO}_{q\bar{q}\to Z^{(*)}}}$$

M_{Z^*} >195 GeV

	$\sigma^{NLO}_{q\bar{q}\to Z^*}$	$\sigma^{NLO}_{q\bar{q}\to ZW}$	$\frac{\sigma_{ZW}}{\sigma_Z}\cdot 10^3$	$\frac{\sigma_{ZW}}{\sigma_{Z^*}}$
Nominal	1898.4	92.5	0.118	0.0487
Maximum	4.6	12.9	16.0	7.9
Minimum	-9.2	-12.0	-11.3	-6.5

Nominal Values of ZZ/Z*

- Ratios are constructed such that the invariant mass of Z* and ZZ are in the same bin
  - Contribution from gg->ZZ increases sigma by ~13%
  - Ratio depends weakly with Mass (nice surprise!)
    - Need to understand better behaviour at very large masses

Cross-sections in fb

Mass Range	$\sigma^{NLO}_{q\bar{q}\to Z^*}$	$\sigma^{NLO}_{q\bar{q}\to ZZ}$	$\sigma^{LO}_{gg\to ZZ}$	$\frac{\sigma_{ZZ}}{\sigma_{Z^*}} \times 10^3$
200 - 250	1773.7	7.99	1.182	5.17
250 - 300	753.2	3.65	0.530	5.54
300 - 350	372.4	1.86	0.246	5.66
350 - 400	205.7	1.07	0.131	5.83
400 - 450	121.0	0.64	0.082	5.94
450 - 500	76.0	0.40	0.055	6.01
500 - 750	143.9	0.74	0.114	5.92
750 - 1000	27.4	0.16	0.033	6.88

Ratio WW/Z$^{(*)}$

- Scale-related uncertainties arise from changing scales by factors of 4 (*4,/4)
  - Pick biggest deviation of changing at the same time and in opposite directions

M_{Z^*} >185 GeV

	$\sigma^{NLO}_{q\bar{q}\to Z}$	$\sigma^{NLO}_{q\bar{q}\to Z^*}$	$\sigma^{NLO}_{q\bar{q}\to WW}$	$\sigma^{LO}_{gg\to WW}$	$\frac{\sigma_{WW}}{\sigma_Z}\cdot 10^3$	$\frac{\sigma_{WW}}{\sigma_{Z^*}}$
Nom.	785.3	2256	636.0	31.04	0.85	0.296
Max.	6.2	4.6	11.5	62.1	16.1	9.4
Min.	-15.7	-9.9	-13.4	-36.0	-8.6	-5.3

Same as above after multiplying σ(gg->WW) by two

	$\sigma^{NLO}_{q\bar{q}\to Z}$	$\sigma^{NLO}_{q\bar{q}\to Z^*}$	$\sigma^{NLO}_{q\bar{q}\to WW}$	$\sigma^{LO}_{gg\to WW}$	$\frac{\sigma_{WW}}{\sigma_Z}\cdot 10^3$	$\frac{\sigma_{WW}}{\sigma_{Z^*}}$
Nom.	785.3	2256.4	636.0	62.08	0.89	0.309
Max.	6.2	4.6	11.5	62.1	19.2	12.0
Min.	-15.7	-9.9	-13.4	-36.0	-10.6	-6.7

Summary

- We have generated MadGraph samples for 4 different topology sets related to same-sign dilepton and multi-lepton signatures, building maps between the theoretical models and the detector observation (PGS) world

- The most important SM backgrounds to each topology are illustrated, though expected to be small. Fitting likelihood is formulated and signals are fitted to pseudodata for each topology. The fits are based on Poisson statistics and are mostly decoupled from the model parameter (shape) assumptions. Multi-dim. fits may be pursued with weakly correlated discriminating variables in the future

- Side band events and control samples (<u>subsidiary measurement</u>) are prepared and formulated into the likelihood to minimize the reliance on the MC prediction for backgrounds. Preliminary exclusion limits are made in terms of σxBR vs. luminosity for each topology. Sensitivity to NP with SS/multi-lep. is promising in ≤1 fb^{-1} of data

- We are getting continued help from the LHC New Physics Working Group, and will investigate more topological signatures and/or construct mass grids, and propogate the methods/results to the ATLAS community

Backup slides

Techinical details about sample generation

- Signal samples and some amount of events for ttbar, W+jets and Z+jets are generated with MadGraph and parton-showered/hadronized with Pythia using the interface ME2pythia.f in the pythia-pgs package

- The W/Z+jets are generated with MLM matching (IEXCFILE=1) for lower parton multiplicities, with showerKt=T and qcut=15, and the highest multiplicity one with IEXCFILE=0. For the W samples, we start from W+2jets as the W inclusive cross section is huge

- Set IMSS(22)=24 in the signal pythia_card.dat file to let Pythia read in the decay table calculated by the scripts from the theorists. Leptonic W decays can be forced in the Pythia cards by MDME commands. For the Squark-pair topology, the supar particle decays (2-body or 3-body) is by BRIDGE directly on the LHE text files (*Veronica Sanz*), and then handed over to Pythia. All signal samples are inclusively produced (no matching) to allow Pythia ISR/FSR

- Once the *_pythia_events.hep.gz files are generated by pythia-pgs, we convert it to root files by ExRootSTDHEPConverter in the ExRootAnalysis package, in which all truth particles (TRootGenParticle) at hadron level are available. We then run the CDF cone4 MidPoint algo. (libs already exist in ExRootAnalysis/tmp/CDFCones/, code available upon request) on the truth particles to form jets (TRootGenJet). The MET is calculated by summing over the p_T of neutrinos and N_1 in the final state

Techinical details about sample generation

- For the signal samples, we run PGS simulation on the *_pythia_events.hep.gz file, specifying cone algorithm and set calorimeter cluster size to 0.4 in the PGS card, to obtain the *_pgs_events.root files readable by ExRootAnalysis. For the background, we simply smear the lepton p_T and jet energy with their expected resolutions. The background statistics we generated is not good, so we did a smoothing of the relevant distributions and threw toys

- As a result, the background “shapes” and “rate” are rough estimates only and are more meant for demonstration purpose. The methods presents in this talk, however, do not rely on these shapes and rates very much. We will perhaps produce more background files that are simulated with PGS (please contact us if you are also interested in these efforts) for, e.g., future workshops