Fake ratios for a same sign double lepton SUSY search

Alina Mreńca

7.09.2011

AGH University of Science and Technology, Cracow

Supervisors: Isabell Melzer-Pellmann, Elías Ron, Matthias Stein

Abstract

This report is the summary of my work in DESY during Summer Student Programme 2011. Fake rate, or tight-to-loose method, is a data-driven method to search for Supersymmetry in a same-sign dileptonic channel. It is applied to predict the Standart Model background which can be observed for certain selection, in this case supersymmetric selection. By subtracting the background from data events, one can determine if there is any excess of data that can be treated as evidence of Supersymmetry existence.

Akcknowledgements

I would like to thank my supervisor Isabell Melzer-Pellmann and everyone who supervised my work during the Summer Student Programme 2011. I would like to express my gratitude to Elías Ron and Matthias Stein for their patience, the time they devoted to me and all their support, as well as to all people I didn't mention, but whose help was very valuable to me.

Contents

1 Introduction

Supersymmetry (SUSY) is symmetry that relates the Standart Model (SM) particles to supersymmetric partner particles that differ by a half unit of spin. For each boson exists a fermion partner and vice versa. A number that is usually conserved in SUSY processes, is **R-parity**, which is defined as:

$$R = (-1)^{3(B-L)+2S}$$

with spin S, baryon number B and lepton number L. If R-parity is preserved,the lightest supersymmetric particle (LSP) cannot decay. It is a very good candidate for missing mass in universe, dark matter.
One of the SUSY channels are final states that contain two leptons with the same charge. A typical SUSY cascade that can result in same-sign (SS) di-lepton final state is quark-quark interaction mediated by gluino, which is the superpartner of gluon. In the chain of decays, more SUSY particles are produced, until reaching lightest SUSY particle (LSP), that doesn't decay due to R-parity conservation.
Since the LSP doesn't interact in the detector, it leaves undetected, producing a large amount of missing energy.

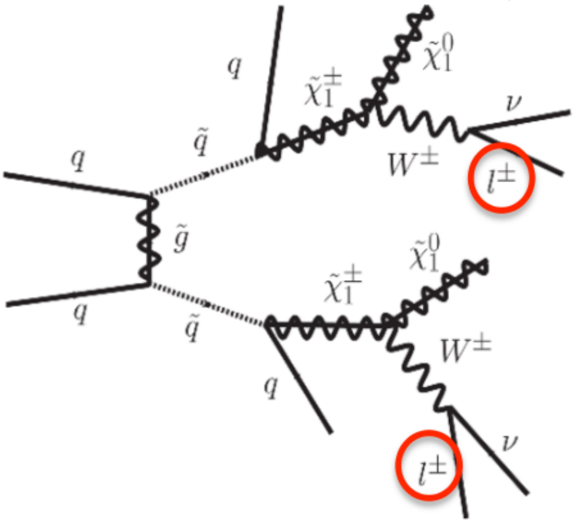


Figure 1: Squark pair production. Prompt leptons are marked with red circles.

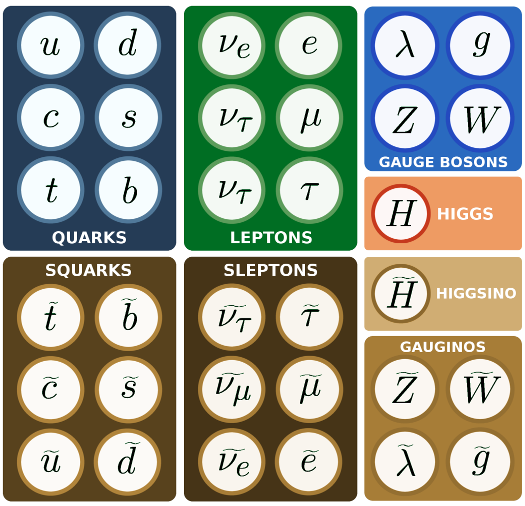


Figure 2: Each particle is related to a particle with spin differing by $\pm\frac{1}{2}$. Superpartners of leptons are sleptons (selectron, squark); names of superpartners of bosons end with "-ino" (gluino, gaugino)

In the Standart Model, processes which result in SS states are very rare. The main backgrounds are processes with one or two fake leptons. These are W+jets and $t\bar{t}$ production with one fake and one prompt lepton, and QCD processes with twi fake leptons: it can happen that both are misidentified as prompt ones. To estimate the rate of these backgrounds, the fake ratio method can be used.

2 Selection & validation plots

Each second up to 40 millions events take place in LHC. The amount of data in CMS is enourmous. A trigger system rejects events from low-energy collisions. It selects potentially interesting events which are not coming from low energy protons collisions. The amount of data is still huge, so some additional and more sophisticated selections must be performed. For example, the most important cuts applied in SUSY selection are:

- Exactly two SS leptons
- $MET > 100\,GeV$ $(MET = |\sum_i \vec{p}_{T,i}|)$ (because LSP leaves undetected)
- $H_T > 250\,GeV$ $(H_T = \sum |\vec{jet}_{p_T}|)$

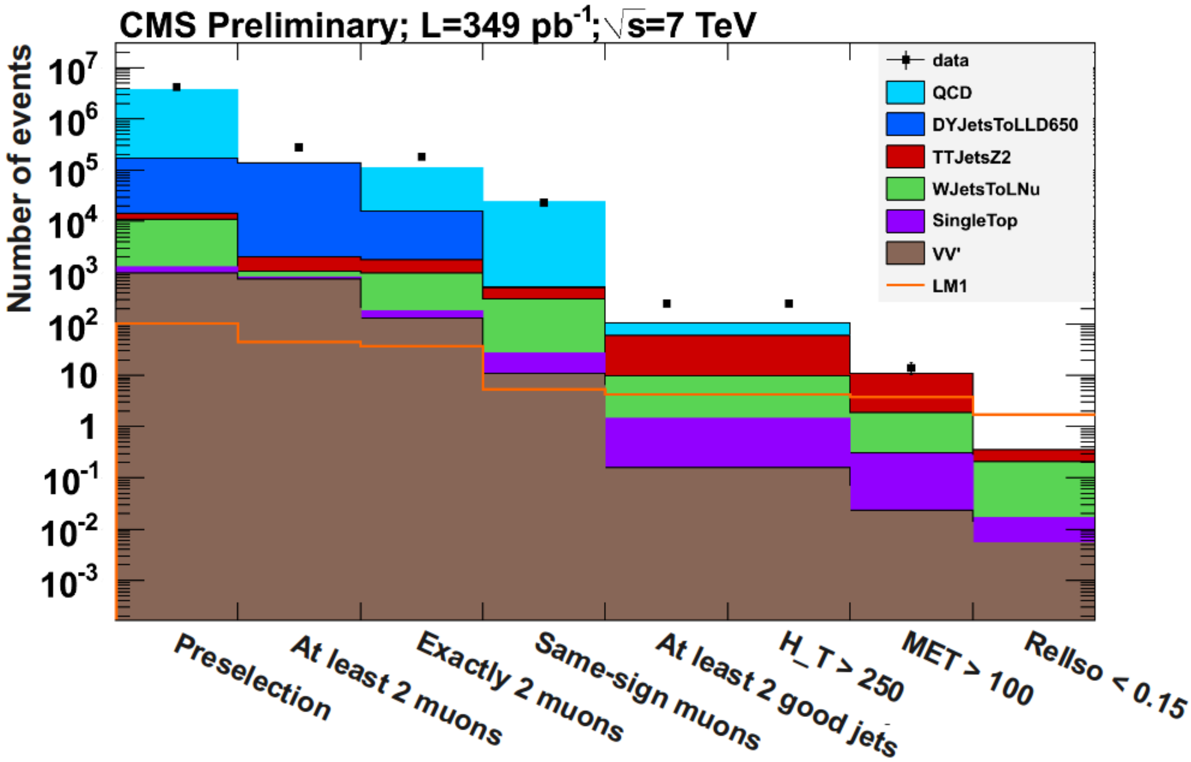


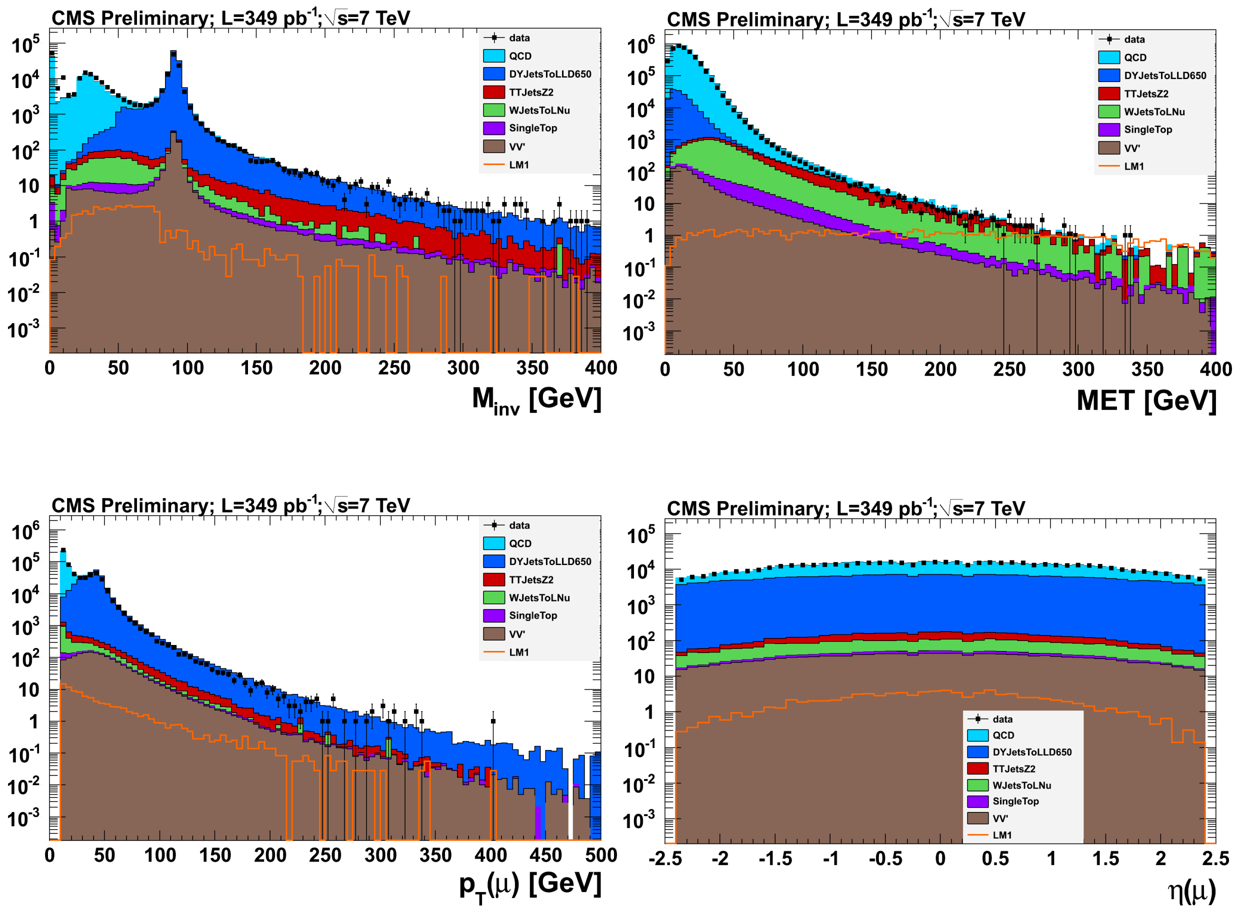
Figure 3: Cut flow for the SUSY selection.

Figure 3 shows the cut flow after the SUSY selection is shown. The number of events surviving each subsequent cut decreases about six orders of magnitude, and finally about 1 per 1 million events survives all the cuts. In comparison, the SUSY scenario known as LM1 (orange line on plots) decreases slowly, and after the selection it is well separated from SM processes.
The validation plots after the preselection are shown in Fig. 4. First plot shows the invariant mass (M_{inv}) between muon pairs, where invariant mass is defined as:

$$M_{inv}(\mu_1\mu_2) = \sqrt{E^2 - |\vec{p}_{\mu_1\mu_2}|^2}$$

$$p_{\mu_1\mu_2} = p_{\mu_1} + p_{\mu_2}$$

The sharp peak corresponds to Z boson, which decays into lepton pairs. Its mass equals 91 GeV, as can be seen on the plot. The second plot shows the missing transverse energy. The following plots show p_L and η distributions for muons and jets. One can see that the agreement between Monte Carlo simulations and real data is relatively good.



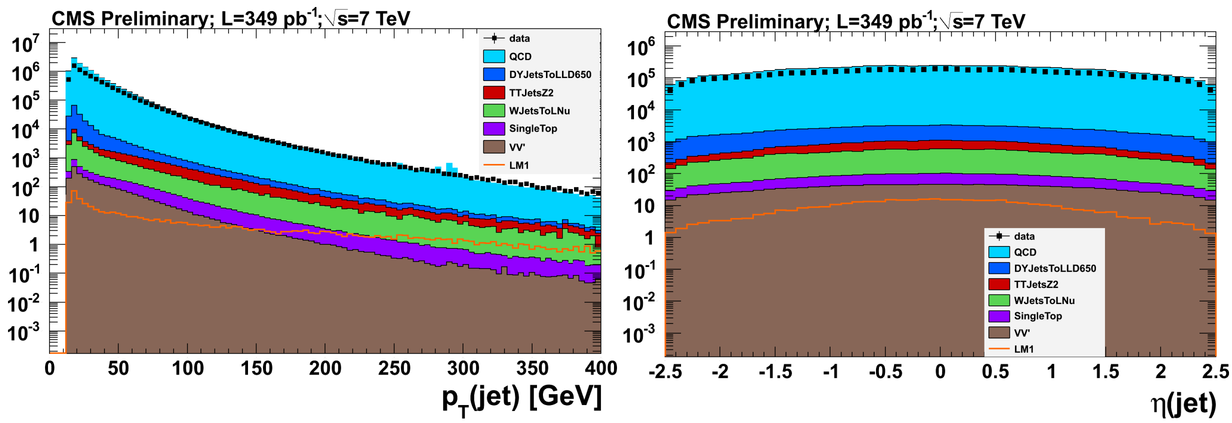


Figure 4: Validation plots.

3 Methodology

In same-sign dileptonic channel, main source of contamination is $t\bar{t}$ process/production. "Fake" electrons are electrons coming from hadronic decays, whereas "prompt" ones come from leptonic W or Z decay, but also from SS chargino decay (SUSY). Both types of leptons can be distinguished by their isolation (Iso) and identification (ID).
Isolation is very important in SUSY searches. Relative isolation (RelIso) is defined as:

$$RelIso = (TrackIso + EcalIso + HcalIso)/p_T$$

It quantifies the activity of jets surronding the lepton and is measured relatively to lepton p_T.
Two sets of leptons are defined:

- loose set, containing of leptons which satisfy loose requirements: $RelIso < 1.0$
- tight leptons, with the tighter isolation requirements $RelIso < 0.15$

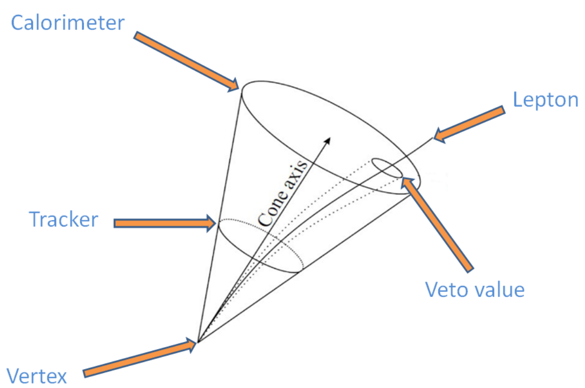


Figure 5: Isolation is used to estimate the activity of jets surrounding the lepton.

All other requirements, like p_T, η cuts, are the same for loose and tight selections.
The "fake ratio" f is the ratio of number of fake leptons passing tight cuts N_t to the number of the leptons which satisfy loose criteria N_l. It can be interpreted as the probability of a fake lepton which is loose to also satisfy tight criteria. The "prompt ratio" of prompt leptons, which pass tight cuts to loose leptons, is a probability for

prompt loose leptons to pass also tight requirements. These ratios are functions of lepton kinematics. It is assumed that they are "universal" and once estimated, can be applied to other backgrounds to extract the signal.
Using f and p and the measurable quantities N_t and N_l, one can derive formulas for the quantities which cannot be directly measured: the number of prompt and fake leptons.

3.1 Fake rates for single lepton events

To better undestand the method, let us start with a simple case of the single lepton channel. Let us assume that:

- p - prompt ratio; probability that a prompt loose lepton is also tight (≈ 1)
- f - fake ratio; probability that a fake loose lepton is also tight (≈ 0)
- N_l - total number of leptons passing loose criteria
- N_p - number of prompt leptons
- N_f - number of fake leptons
- N_{t0} - number of events where no lepton passes tight criteria
- N_{t1} - number of events where one lepton passes tight criteria

All of these quantities depend on lepton's kinematics (p_T and η).
N_p and N_f are related to the number of events where 0 or 1 leptons pass tight cuts by:

$$N_{t0} = (1-p)N_p + (1-f)N_f$$
$$N_{t1} = pN_p + fN_f$$

The interpretation of these relations is as follows. Number of events with one tight lepton N_{t1} contains prompt leptons which pass tight criteria pN_p, and fake leptons which are counted as tight ones fN_f. Similarly, N_{t0} consist of prompt leptons that don't pass tight cuts $(1-p)N_p$ and fake leptons which are counted as fake $(1-f)N_f$.
By inverting these relations one obtains:

$$N_p = \frac{1}{p-f}[(1-f)N_{t1} - fN_{t0}]$$
$$N_f = \frac{1}{p-f}[pN_{t0} - (1-p)N_{t1}]$$

Using these quantities one can derive the number of prompt signal events

$$N_{signal} = N_P^{pass} = pN_p$$

and number of fake passing the tight cuts (contamination)

$$N_{contam} = N_f^{pass} = fN_f$$

3.2 Fake rates for dileptonic events

Same flavour events is relevant $t\bar{t}$ processes and in SUSY for SS dileptons. Two leptons are distinguished by their momenta. The notation is N_{xy}, $x, y = p, f$, where the first label refers to leptons with higher p_T and the second label to the lepton with lower p_T. Let us assume that:

- N_{pp} - number of events with two prompt leptons
- N_{pf} - number of events with one prompt and one fake lepton
- N_{ff} - number of events with two fake leptons
- N_{txy} (x,y=0,1) - number of events with 0,1,or 2 leptons passing tight cuts

Where N_{xy}, N_{txy} are functions of leptons' p_T and η. Basic relations between them:

$$N_{t00} = (1-p_1)(1-p_2)N_{pp} + (1-p_1)(1-f_2)N_{pf} + (1-f_1)(1-p_2)N_{fp} + (1-f_1)(1-f_2)N_{ff}$$

$$N_{t10} = p_1(1-p_2)N_{pp} + p_1(1-f_2)N_{pf} + f_1(1-p_2)N_{fp} + f_1(1-f_2)N_{ff}$$

$$N_{t01} = (1-p_1)p_2N_{pp} + (1-p_1)f_2N_{pf} + (1-f_1)p_2N_{fp} + (1-f_1)f_2N_{ff}$$

$$N_{t11} = p_1p_2N_{pp} + p_1f_2N_{pf} + f_1p_2N_{fp} + f_1f_2N_{ff}$$

From these equations one can derive:

$$A = \begin{pmatrix} f_1f_2 & -(1-f_1)f_2 & -f_1(1-f_2) & (1-f_1)(1-f_2) \\ -f_1p_2 & (1-f_1)p_2 & f_1(1-p_2) & -(1-f_1)(1-p_2) \\ -p_1f_2 & (1-p_1)f_2 & p_1(1-f_2) & -(1-p_1)(1-f_2) \\ p_1p_2 & -(1-p_1)p_2 & -p_1(1-p_2) & (1-p_1)(1-p_2) \end{pmatrix}$$

$$\begin{pmatrix} N_{pp} \\ N_{pf} \\ N_{fp} \\ N_{ff} \end{pmatrix} = \frac{A}{(p_1-f_1)(p_2-f_2)} \begin{pmatrix} N_{t00} \\ N_{t01} \\ N_{t10} \\ N_{t11} \end{pmatrix}$$

The contributions to backgrounds from the selected sample are

- $(p_1f_2N_{pf} + p_2f_1N_{fp})$ - from W + jets events (one fake lepton)
- $f_1f_2N_{ff}$ - from QCD (two fake leptons)

In same sign SUSY events, main background contributing to N_{fp} and N_{pf} are $t\bar{t}$ with one prompt lepton from the W decay and fake from the leptonic decay of b-quark from coming from top. The N_{ff} part comes from QCD background.

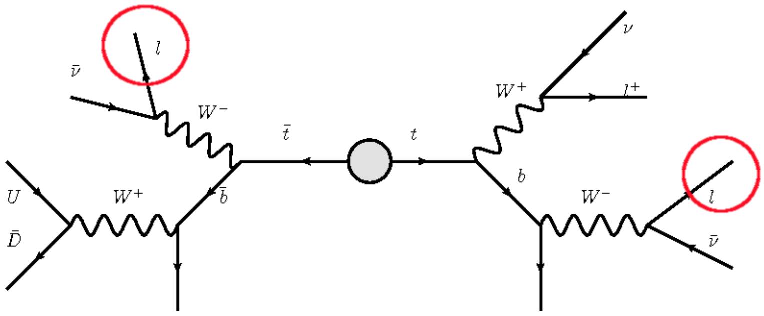


Figure 6: $t\bar{t}$ decay resulting in two same-sign leptons final state. The prompt lepton comes from the W decay and the fake one from b quark decay.

4 Results

The prompt ratio can be calculated in events with Z + jets, (using tag-and-probe method). To measure the prompt ratio, events with exactly two leptons are considered. Let us define histograms with bins in p_T, which contain:

- N_{TT_h} - number of leptons with higher p_T (event with two tight leptons)
- N_{TT_l} - number of leptons with lower p_T (event with two tight leptons)
- N_{LT} - number of loose leptons if loose is the one with higher p_T (event with one tight lepton)
- N_{TL} - number of loose leptons if loose is the one with lower p_T (event with one tight lepton)

The ratio is defined as

$$r = \frac{N_{TT_h} + N_{TT_l}}{N_{TT_h} + N_{TT_l} + N_{LT} + N_{TL}}$$

The ratio always has a value between 0 and 1, since the numerator is always a subset of denominator.

4.1 Prompt ratio measurements

We expect the prompt ratio to be close to 1, since for the prompt muon the probability that it is also isolated is high.

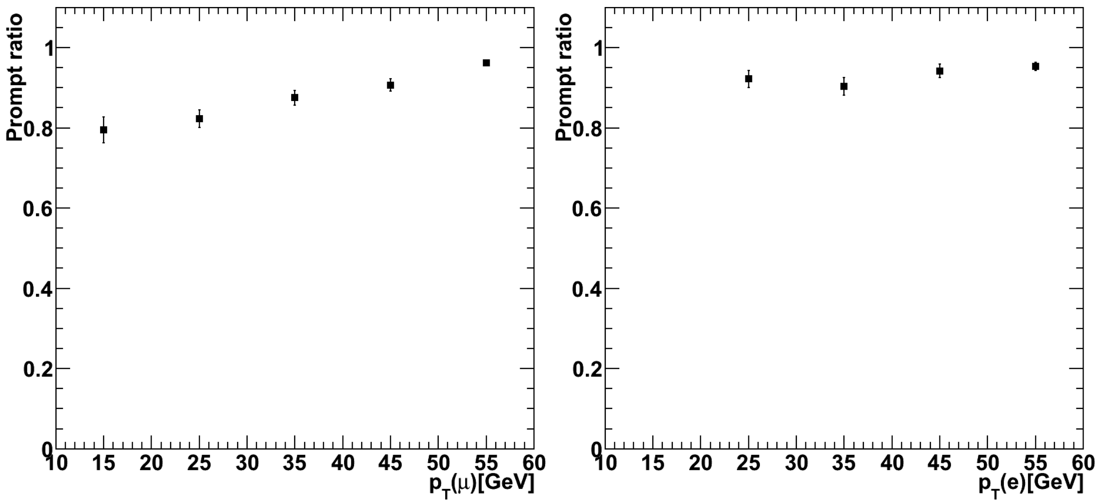


Figure 7: Prompt ratio for muons (left hand side) and electrons (right hand side) as function of p_T

The event selection applied for prompt ratio favours muons from Z decays:

- Exactly two loose muons
- At least one tight muon
- Invariant mass between the two muons inside 15 GeV of Z boson mass, opposite charge
- Missing transverse energy $<$ 20 GeV
- At least two good jets ($p_T > 40$ GeV)

For electron prompt ratio following selection was used:

- Exactly two loose electrons
- At least one tight electron
- Invariant mass between the two electrons inside 15 GeV of the Z boson mass, opposite charge
- Missing transverse energy $<$ 20 GeV
- At least two good jets ($p_T > 40$ GeV)

4.2 Measurements for electrons

On the other hand, the fake ratio is expected to take the values close to 0, as the probability for a loose lepton to pass tight cuts is small.

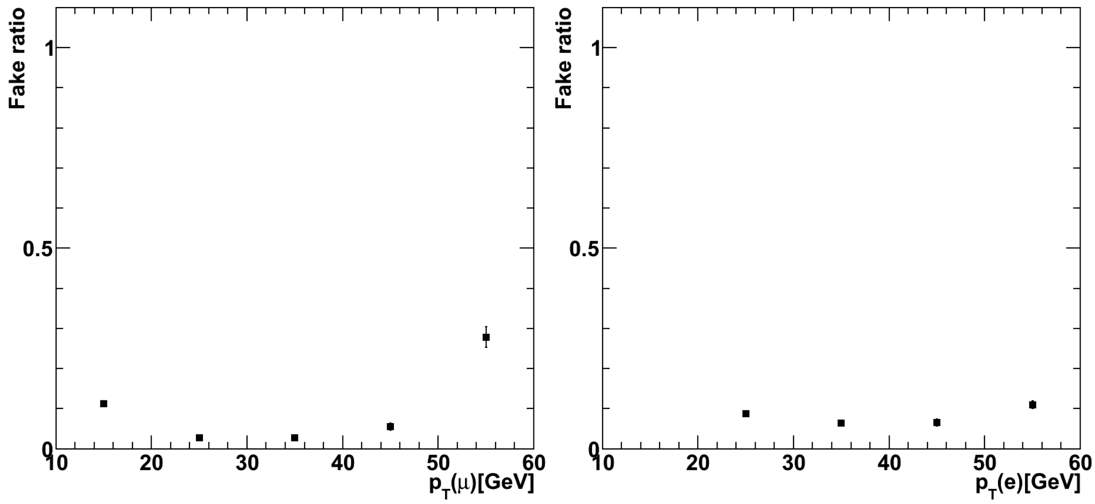


Figure 8: Fake ratio for muons (left hand side) and electrons (right hand side) as function of p_T

To measure the fake ratio for electrons a selection suppressing the contribution of prompt electrons from Z and W decays:

- Exactly one loose electron
- Transverse energy between muon and MET < 20 GeV
- Missing transverse energy < 20 GeV
- At least two good jets ($p_T > 40$ GeV)

For electrons, following selection was designed:

- Exactly one electron passing loose cuts
- Tranverse mass between electron and MET < 20 GeV
- Missing transverse energy < 20 GeV
- At least two good jets ($p_T > 40$ GeV)

The number of events measured after the SUSY selection in Monte Carlo, real data, LM1 model and predicted by the tight-to-loose method covered in this report, can be summarized in the Table 1. What we can learn from it, is that the observed number of events in data is consistent with the one predicted by fake rate method. The yield for the SUSY scenario LM1, shown in the last row, predicts a lot more events than the ones measured. Since no excess of data is observed, there is no evidence of SUSY existence.

	$\mu\mu$	$e\mu$	ee	total
MC	0.95±0.37	1.07±0.33	0.32±0.13	2.34±0.5
Pred. (TL)	1.94±0.57	1.4±0.42	0.39±0.19	3.73±0.73
Data (1.1fb)	1±1	2±1.4	0	3±1.7
Signal (LM1)	13.4±1.3	18.3±1.7	5.6±0.9	37.3±2.3

Table 1: Number of events measured in Monte Carlo, data, LM1 and predicted by fake rate method.

5 Conclusion

To sum up, the tight-to-loose method is a useful tool for estimating the Standard Model background after the SUSY selection. For doing this, the prompt ratio and the fake ratio for muons and electrons were calculated as a function of p_T. The validation on Monte Carlo samplesis successful - we observe a good agreement between MC and data. Also, the cut flow for Supersymmetry selection was established. The quantities obtained through data selection and predicted by fake rate method are compatible, but unfortunately, there is no excess of data seen that can be interpreted as SUSY evidence.

References

[1] H. Bakhshian et al., CMS AN -2010/261

[2] H. Bakhshian et al., CMS AN -2011/262