Update on the progress of the Object Selection

The last results

A. FAHIM[1]

[1]IPM & Sharif University

IPM Weekly Report, Jan 19, 2010

Outline

Outline

The methods used for optimization

There are some methods to optimize the selection cuts and choose the critical region in object selection stage:

- "One way of finding an optimal placement of the cuts is to require that they give a maximum purity for a given efficiency" Cowan; Statistical Data Analysis, page 50
- maximum of Efficiency (1-α) and minimum of Mis-id probability(β) Lista, Part 1.2 section *Cut analysis*[1]
- Optimize the significance $\frac{n_{sig}}{\sqrt{n_{sig}+n_{bkg}}}$ after the overall event selection.

It seems that the last one is the best way with looking at a big multi-dimensional space of all the observables. Hence I follow this stage via second method and optimization has to be postponed until event selection stage.

[1] http://indico.cern.ch/getFile.py/access?contribId=3&resId=4&materialId=slides& (as I learned from slides)

The methods used for optimization

There are some methods to optimize the selection cuts and choose the critical region in object selection stage:

- "One way of finding an optimal placement of the cuts is to require that they give a maximum purity for a given efficiency" Cowan; Statistical Data Analysis, page 50
- maximum of Efficiency (1-α) and minimum of Mis-id probability(β) Lista, Part 1.2 section *Cut analysis*[1]
- Optimize the significance $\frac{n_{sig}}{\sqrt{n_{sig}+n_{bkg}}}$ after the overall event selection.

It seems that the last one is the best way with looking at a big multi-dimensional space of all the observables. Hence I follow this stage via second method and optimization has to be postponed until event selection stage.

[1] http://indico.cern.ch/getFile.py/access?contribId=3&resId=4&materialId=slides& (as I learned from slides)

The methods used for optimization

There are some methods to optimize the selection cuts and choose the critical region in object selection stage:

- "One way of finding an optimal placement of the cuts is to require that they give a maximum purity for a given efficiency" Cowan; Statistical Data Analysis, page 50
- maximum of Efficiency (1-α) and minimum of Mis-id probability(β) Lista, Part 1.2 section *Cut analysis*[1]
- Optimize the significance $\frac{n_{sig}}{\sqrt{n_{sig}+n_{bkg}}}$ after the overall event selection.

It seems that the last one is the best way with looking at a big multi-dimensional space of all the observables. Hence I follow this stage via second method and optimization has to be postponed until event selection stage.

[1] http://indico.cern.ch/getFile.py/access?contribId=3&resId=4&materialId=slides& (as I learned from slides)

The methods used for optimization

There are some methods to optimize the selection cuts and choose the critical region in object selection stage:

- "One way of finding an optimal placement of the cuts is to require that they give a maximum purity for a given efficiency" Cowan; Statistical Data Analysis, page 50
- maximum of Efficiency (1-α) and minimum of Mis-id probability(β) Lista, Part 1.2 section *Cut analysis*[1]
- Optimize the significance $\frac{n_{sig}}{\sqrt{n_{sig}+n_{bkg}}}$ after the overall event selection.

It seems that the last one is the best way with looking at a big multi-dimensional space of all the observables. Hence I follow this stage via second method and optimization has to be postponed until event selection stage.

[1] http://indico.cern.ch/getFile.py/access?contribId=3&resId=4&materialId=slides& (as I learned from slides)

The methods used for optimization

There are some methods to optimize the selection cuts and choose the critical region in object selection stage:

- "One way of finding an optimal placement of the cuts is to require that they give a maximum purity for a given efficiency" Cowan; Statistical Data Analysis, page 50
- maximum of Efficiency (1-α) and minimum of Mis-id probability(β) Lista, Part 1.2 section *Cut analysis*[1]
- Optimize the significance $\frac{n_{sig}}{\sqrt{n_{sig}+n_{bkg}}}$ after the overall event selection.

It seems that the last one is the best way with looking at a big multi-dimensional space of all the observables. Hence I follow this stage via second method and optimization has to be postponed until event selection stage.

[1] http://indico.cern.ch/getFile.py/access?contribId=3&resId=4&materialId=slides& (as I learned from slides)

Outline

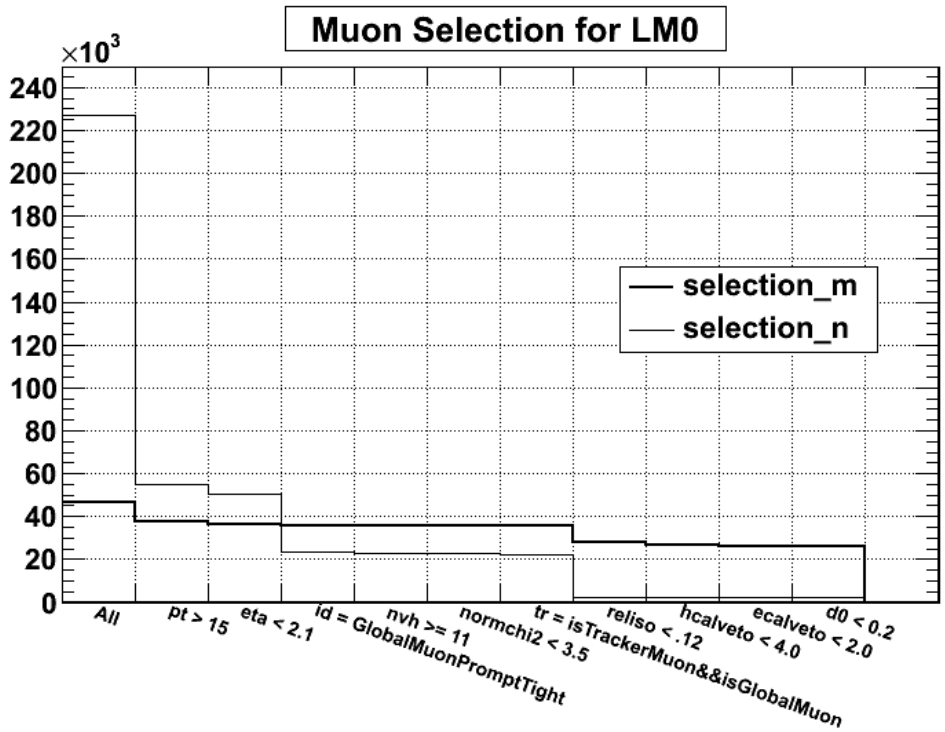
Some Preliminary Information

- CMSSW version is 3_1_2 in both data skimming and analysis
- Data sample is LM0: /LM0/Summer09-MC_31X_V3-v1/GEN-SIM-RECO
- pattuples are produced by pre-selections cuts:
  - process.selectedLayer1Muons.cut = cms.string('pt > 3. & abs(eta) < 3.5')
  - process.selectedLayer1Electrons.cut = cms.string('pt > 3. & abs(eta) < 3.5')
  - process.selectedLayer1Jets.cut = cms.string('et > 3. & abs(eta) < 3.5')

Outline

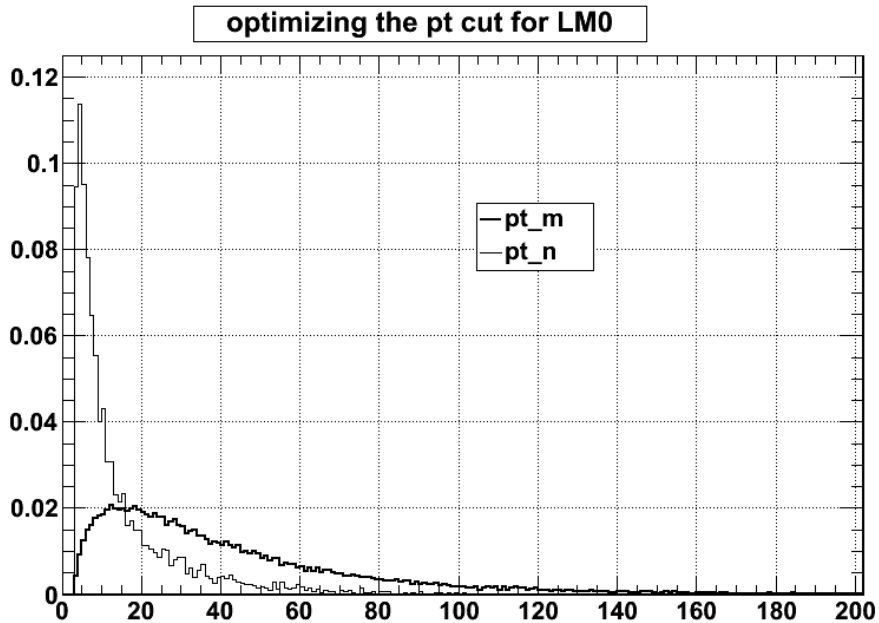
Summary of Cuts

- The cuts are imposed step by step
- normchi2 has no serous effect on selection because GlobalMuonPromptTight Identification includes this cut.
- ecalveto and hcalveto kill more the signal than background! mu->isolationR03().emVetoEt (dR<0.07 in ecal and dR<0.1 in hcal).
- Impact parameter has also no effect.



Pt of Muons

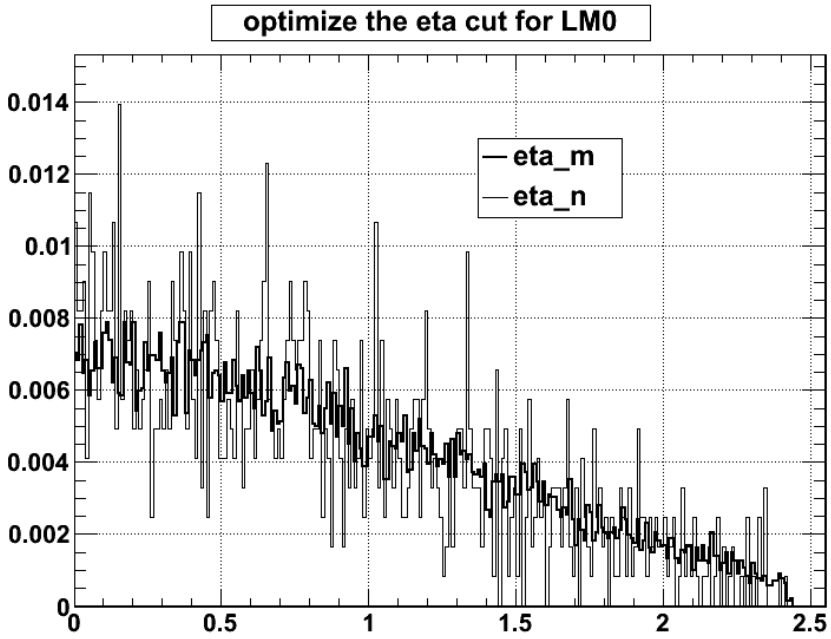
- $|\eta| < 2.1$
- id = isGlobalMuonPromptTight
- Number of Valid Hits of Track ≥ 11
- $\chi^2_{norm} < 3.5$
- track type = isGlobal & isTracker
- RelIso = $\frac{trackerIso+calorimeterIso}{max(20,p_t)} < 0.12$
- hcalveto < 4.0
- ecalveto < 2.0
- $d0_{primaryVertex} < 0.2$



- $p_t > 15$ seems a good selection for muon
- pt_m is related to reconstructed muons which are matched to generated MC muon and pt_n related to ones are not matched

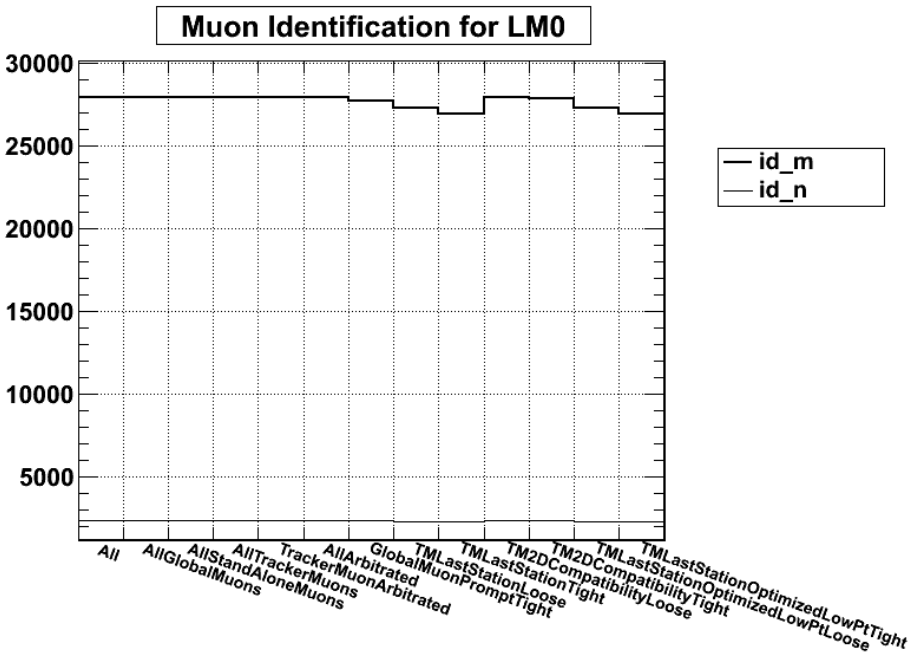
Eta of Muons

- $p_t > 15$
- id = isGlobalMuonPromptTight
- Number of Valid Hits of Track ≥ 11
- $\chi^2_{norm} < 3.5$
- track type = isGlobal & isTracker
- RelIso = $\frac{trackerIso+calorimeterIso}{max(20,p_t)} < 0.12$
- hcalveto < 4.0
- ecalveto < 2.0
- $d0_{primaryVertex} < 0.2$



- $|\eta|$can not discriminate matched muons from non-matched ones.

Identification of Muons

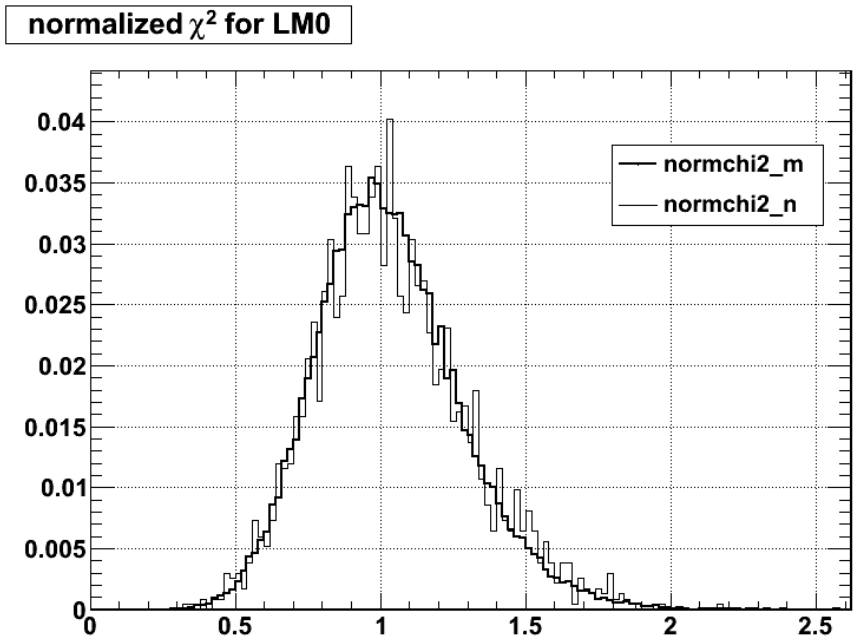


- TM2DCompatibilityLoose/Tight work better than GlobalMuonPromptTight

- $p_t > 15$
- $| \eta | < 2.1$
- Number of Valid Hits of Track ≥ 11
- $\chi^2_{norm} < 3.5$
- track type = isGlobal & isTracker
- RelIso = $\frac{trackerIso+calorimeterIso}{max(20,p_t)} < 0.12$
- hcalveto < 4.0
- ecalveto < 2.0
- $d0_{primaryVertex} < 0.2$

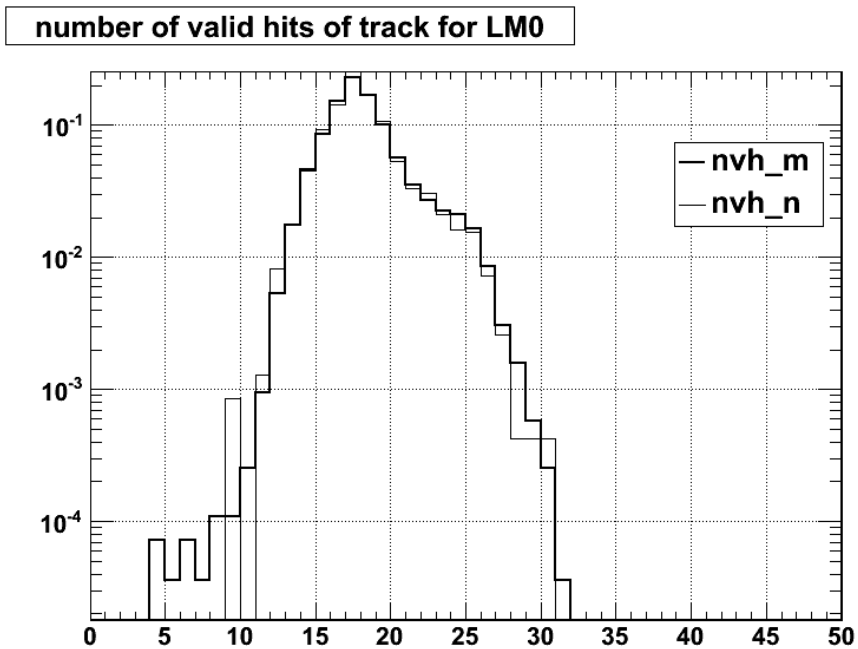
χ^2_{norm} of Muons Track

- $p_t > 15$
- id = isGlobalMuonPromptTight
- Number of Valid Hits of Track ≥ 11
- $|\eta| < 2.1$
- track type = isGlobal & isTracker
- RelIso = $\frac{trackerIso+calorimeterIso}{max(20,p_t)} < 0.12$
- hcalveto < 4.0
- ecalveto < 2.0
- $d0_{PrimaryVertex} < 0.2$



- χ^2_{norm} can not discriminate matched muons from non-matched ones

Number of Valid Hits of Muons Track

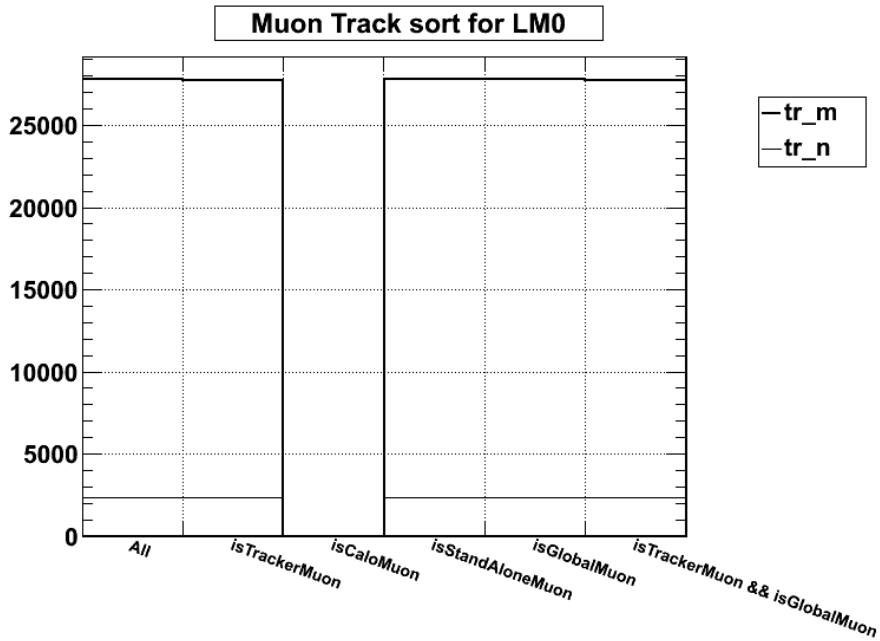


- $p_t > 15$
- $| \eta | < 2.1$
- id = isGlobalMuonPromptTight
- $\chi^2_{norm} < 3.5$
- track type = isGlobal & isTracker
- RelIso = $\frac{trackerIso+calorimeterIso}{max(20,p_t)} < 0.12$
- hcalveto < 4.0
- ecalveto < 2.0
- $d0_{PrimaryVertex} < 0.2$

- TNumber of Valid Hits of Track ≥ 11 is good

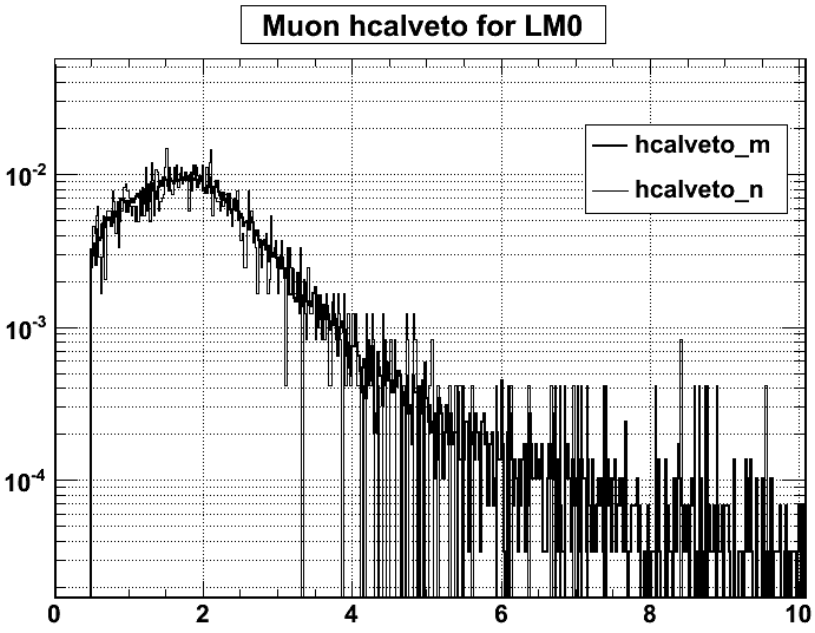
Track Type of Muons

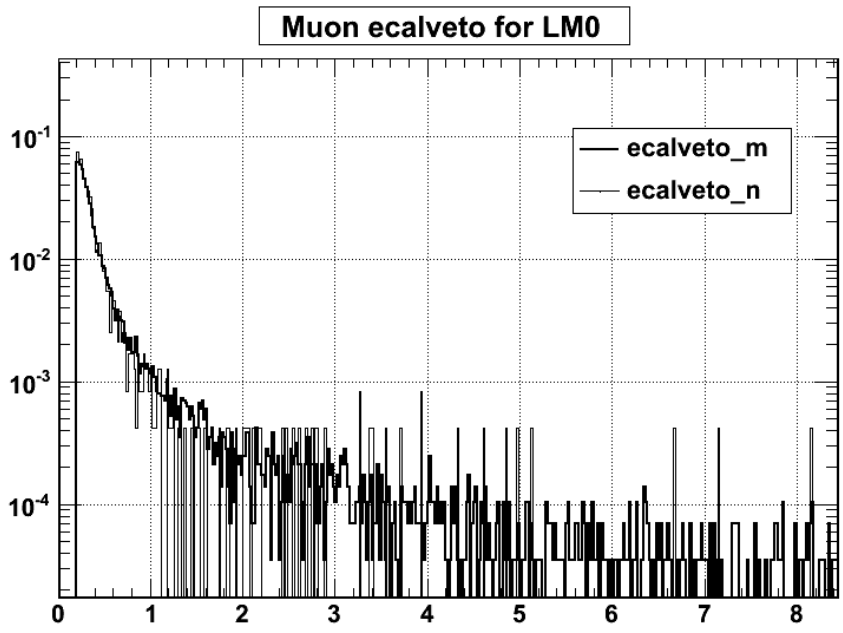
- $p_t > 15$
- id = isGlobalMuonPromptTight
- Number of Valid Hits of Track ≥ 11
- $\mid \eta \mid < 2.1$
- $\chi^2_{norm} < 3.5$
- RelIso = $\frac{trackerIso+calorimeterIso}{max(20,p_t)} < 0.12$
- hcalveto < 4.0
- ecalveto < 2.0
- $d0_{PrimaryVertex} < 0.2$



- All Muons are stanalone and Global and there is not any caloMuon !

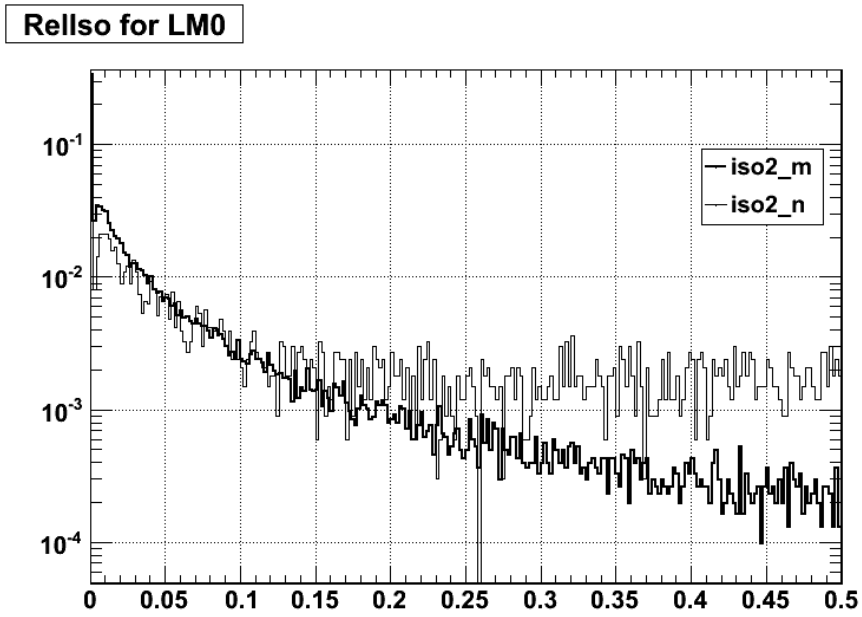
ecal and hcal veto of Muons





- Neither hcalveto nor ecalveto can discriminate matched muons from non-matched ones.

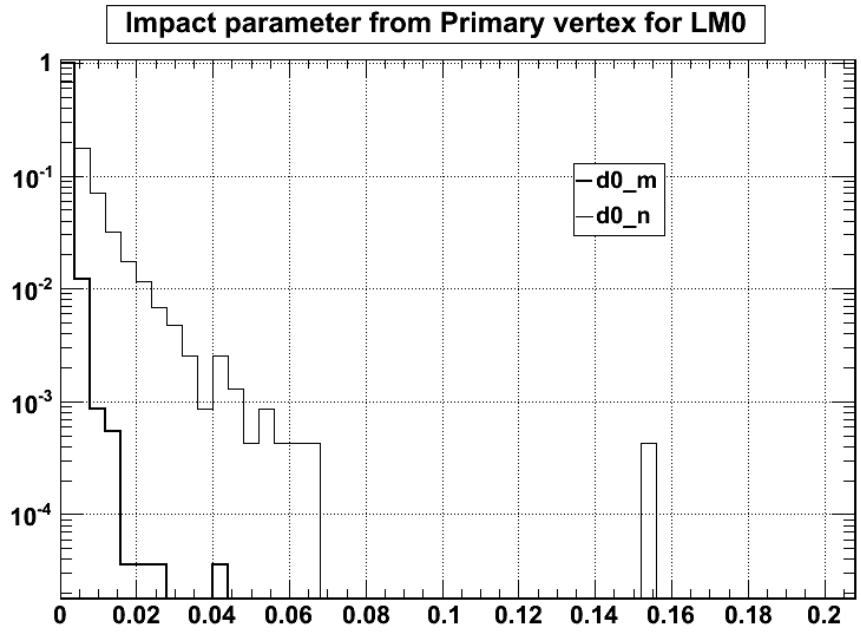
Isolation Parameter of Muons



- RelIso = $\frac{trackerIso+calorimeterIso}{max(20,p_t)} < 0.15$ is a good cut

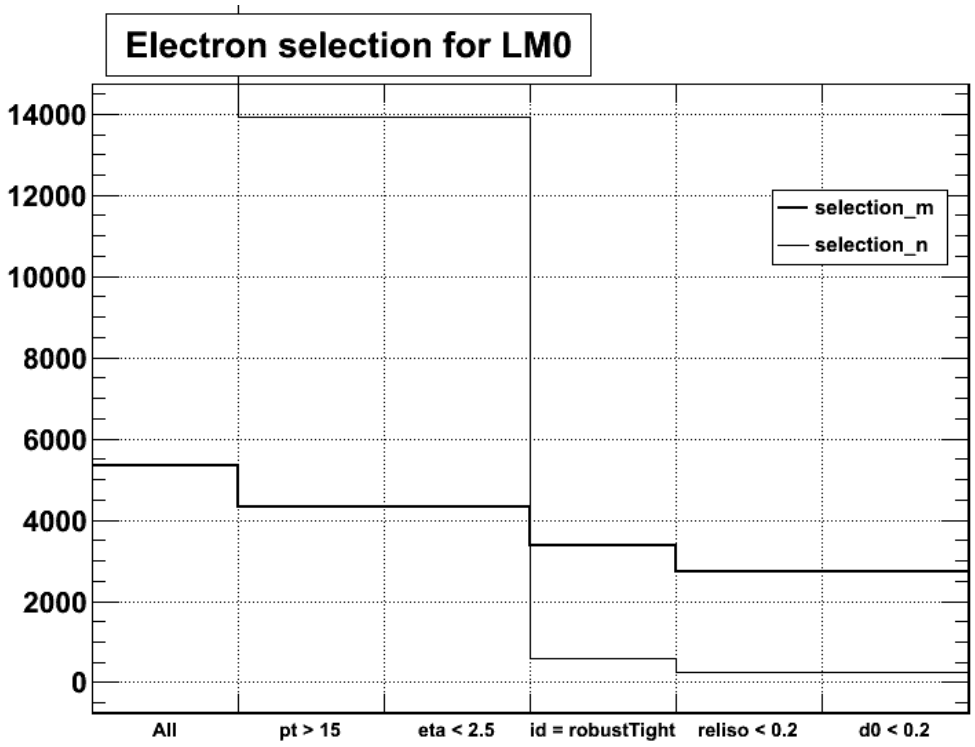
Impact Parameter of Muons

- Impact Parameter from Primary Vertex < 0.02 is good (200μm)



Outline

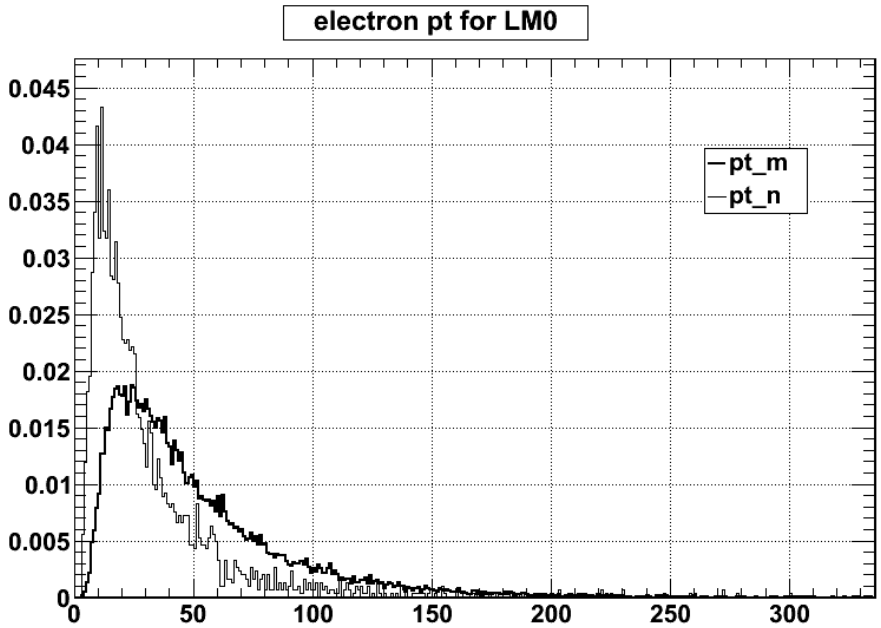
Summary of Cuts



The cuts are imposed step by step

Pt of Electrons

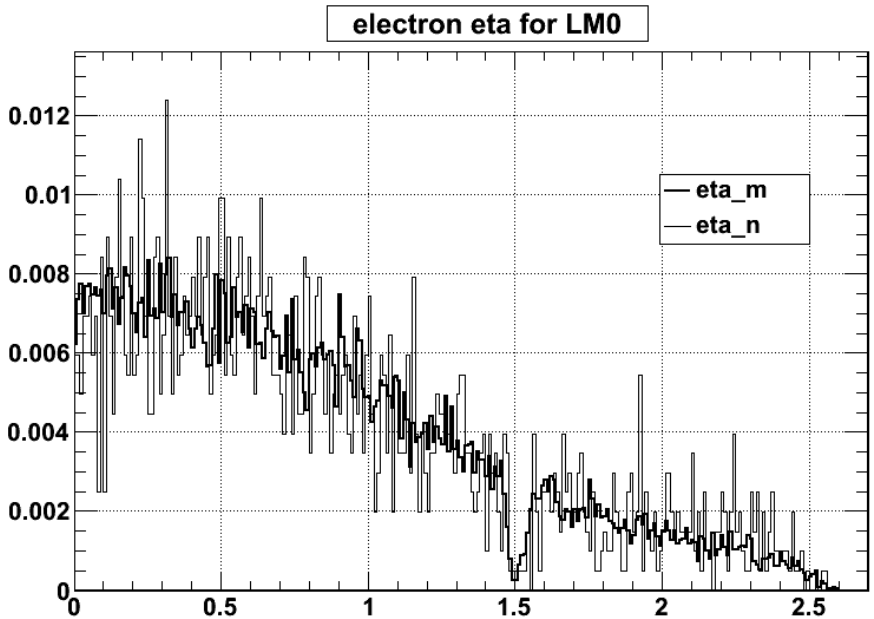
- $|\eta| < 2.5$
- RelIso = $\frac{trackerIso+calorimeterIso}{max(20,p_t)} < 0.2$
- Id = robustTight
- $d0_{PrimaryVertex} < 0.2$



- $p_t > 20$ is better.
- pt_m is related to reconstructed electrons which are matched to generated MC electrons and pt_n related to ones are not matched

Eta of Electrons

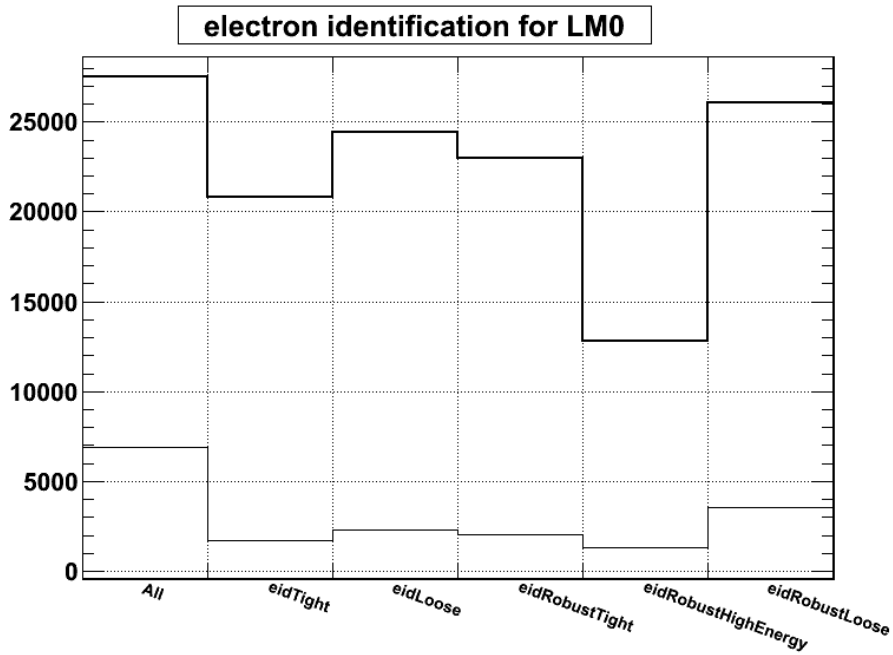
- $p_t > 15$
- RelIso = $\frac{trackerIso+calorimeterIso}{max(20,p_t)} < 0.2$
- Id = robustTight
- $d0_{PrimaryVertex} < 0.2$



- $|\eta|$is not a good discriminator

Identification of Electrons

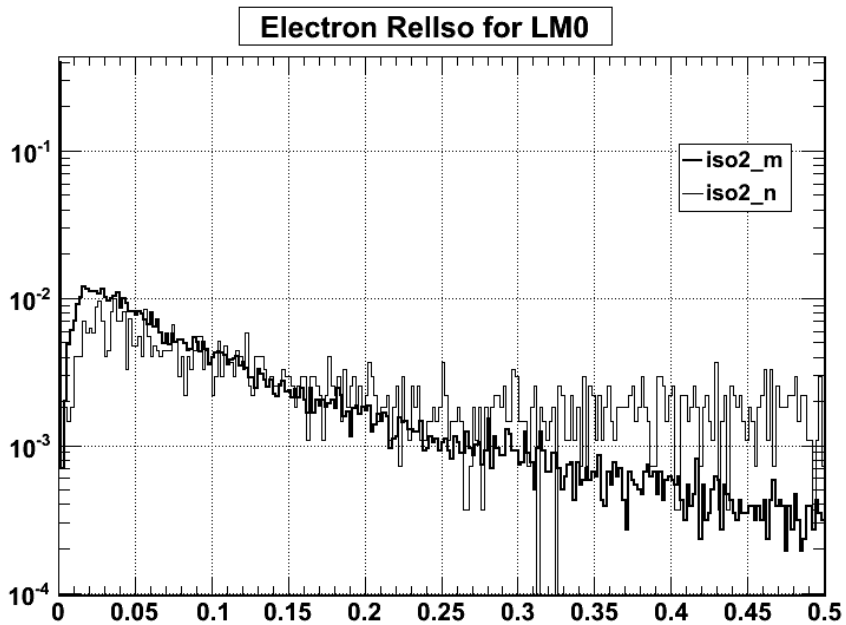
- $p_t > 15$
- RelIso = $\frac{trackerIso+calorimeterIso}{max(20,p_t)} < 0.2$
- $| \eta | < 2.5$
- $d0_{PrimaryVertex} < 0.2$



- RobustTight is a good identification

Isolation of Electrons

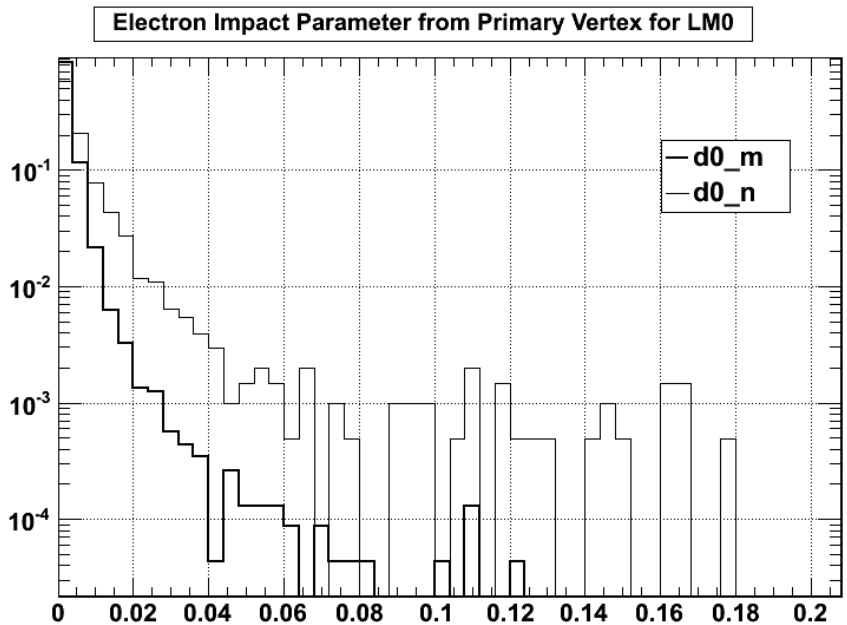
- $p_t > 15$
- Id = robustTight
- $| \eta | < 2.5$
- $d0_{PrimaryVertex} < 0.2$



- RelIso = $\frac{trackerIso+calorimeterIso}{max(20,p_t)} < 0.2$

Impact Parameter from Primary Vertex of Electrons

- $p_t > 15$
- Id = robustTight
- $| \eta | < 2.5$
- RelIso = $\frac{trackerIso+calorimeterIso}{max(20,p_t)} < 0.2$

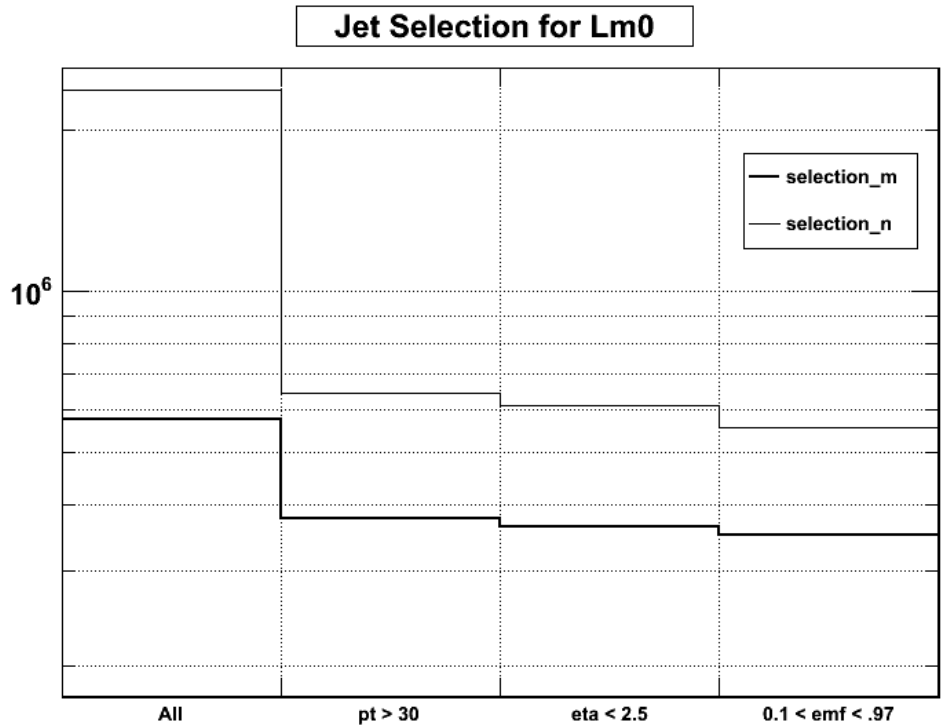


- $d0_{PrimaryVertex} < 0.02$

Outline

Summary of Cuts

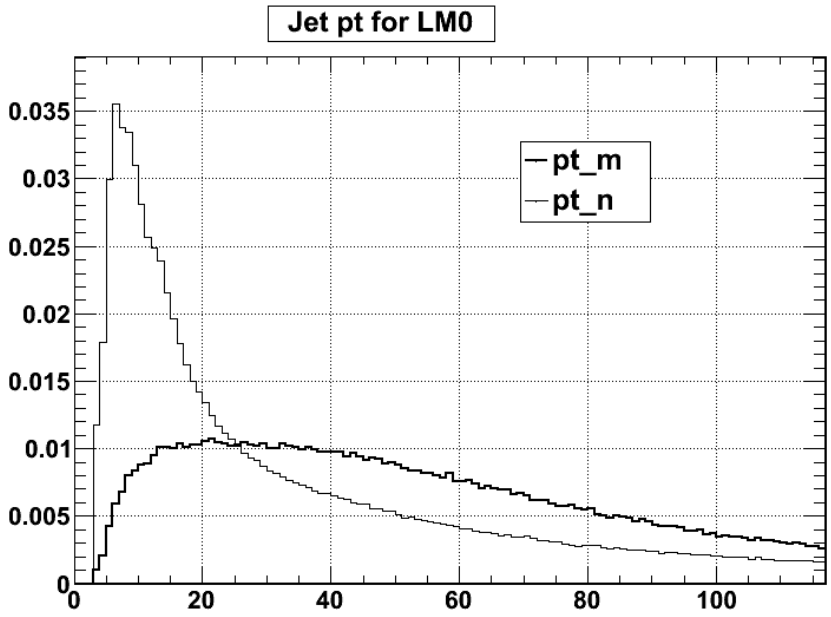
- $p_t > 30$
- $|\eta| < 2.5$
- 0.1 < emf < 0.97



- purity of jet selection is very low and it should get better after cleaning from leptons

Corrected Pt of Jets

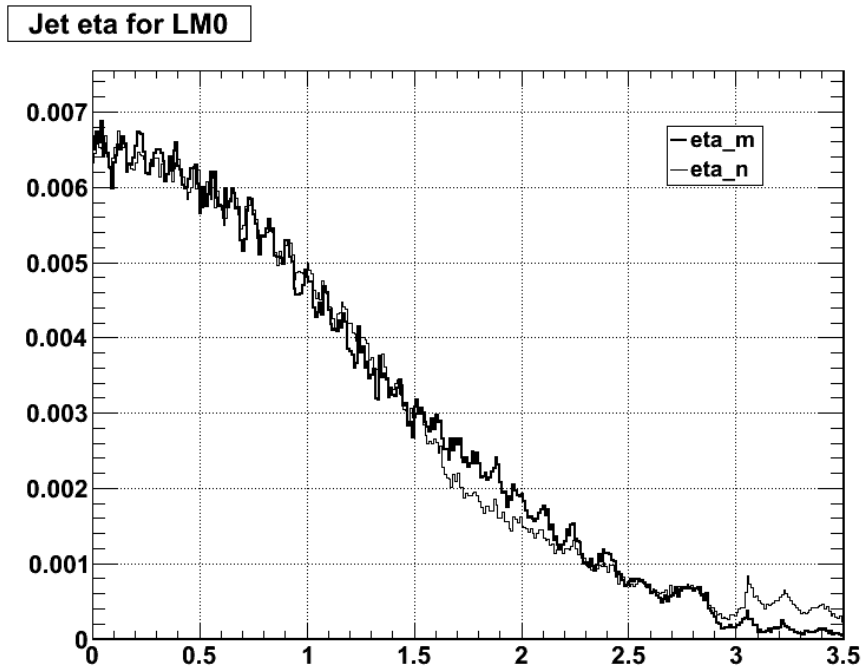
- $|\eta| < 2.5$
- $0.1 < \text{emf} < 0.97$



- $p_t > 30$

Eta of Jets

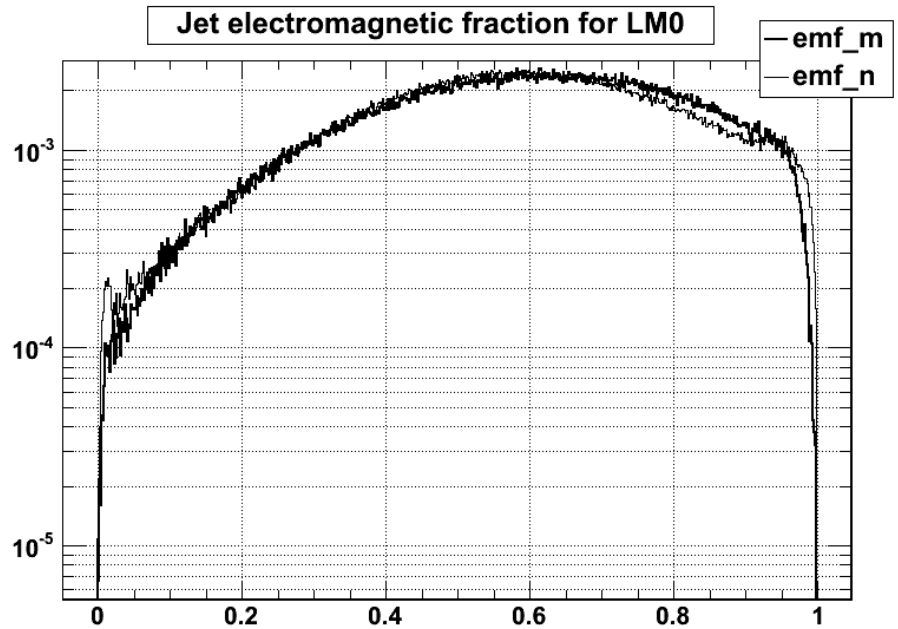
- $p_t > 30$
- $0.1 < \text{emf} < 0.97$



- $|\eta| < 3.$

Electromagnetic Fraction of Jets

- $p_t > 30$
- $|\eta| < 2.5$



- 0.1 < emf < 0.97